DOCUMENT RESUME

ED 216 036 TM 820 198

AUTHOR Wisenbaker, Joseph M.; Kolstad, Andrew J.
TITLE Factors Related to Third Follow-up Survey Responses: Contractor Report.
INSTITUTION Research Triangle Inst., Durham, N.C. Center for Educational Research and Evaluation.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-82-209
PUB DATE Sep 81
CONTRACT OEC-0-73-6666
NOTE 80p.; For related document, see TM 820 197; small print in tables.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Correlation; *Data Collection; Grade 12; *Graduate Surveys; High Schools; *Longitudinal Studies; Outcomes of Education; Responses; *Student Characteristics; Vocational Followup
IDENTIFIERS *National Longitudinal Study High School Class 1972; *Response Rates (Questionnaires)

ABSTRACT

A third followup survey within a National Longitudinal Study (NLS) of the high school class of 1972 finds that there is little difference in group variability in survey response when considering background characteristics, while significant differences in response rates are found when considering the group patterns of response to prior instruments. Prior response is given as the most important predictor of third followup response. Meant to measure postsecondary educational and vocational activities, the survey is intended to guide federal planning and implemention of educational and career programs. The scope and purpose and descriptive statistics are presented analyzing response rates, response modes and their quality. The prediction approaches and results of the variables; a discriminate function analysis of mail and personal interview modes; and the quality of data are discussed, referring to the importance of a response payment incentive procedure. The study is meant to facilitate future data collection efforts. Appendices and tables illustrate the analysis. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216036

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Contractor Report

Factors Related to Third Follow-Up Survey Responses

National Center for
Education Statistics

TM 820 198

Factors Related to Third Follow-Up Survey Responses

Center for Educational Research and Evaluation

Joseph M. Wisenbaker

Andrew J. Kolstad
Project Officer
National Center for Education Statistics

September 1981

Prepared for the National Center for Education Statistics under contract OE-0-73-6666 with the U.S. Department of Education. Contractors undertaking such projects are encouraged to express freely their professional judgment. This report, therefore, does not necessarily represent positions on policies of the Government, and no official endorsement should be inferred. This report is released as received from the contractor.

NCES 82-209

ACKNOWLEDGMENTS

The authors wish to thank all those people who have contributed to this study and the writing of this report. Grateful acknowledgments are due, in particular, to Dr. Jay Levinsohn, former NLS Project Director, and Ms. Louise Henderson. Acknowledgments are also due to Dr. Andrew Kolstad and Mr. Kenneth Tabler of the National Center for Education Statistics, Dr. Graham Burkheimer, and Dr. John A. Riccobono, NLS Project Director, for their significant input and assistance. Special thanks are due to Mrs. Pam Mikels for her expert assistance in the preparation of this report.

CONTENTS

		Page
	Foreword	iii
	Acknowledgments	iv
I.	INTRODUCTION	1
	A. General Description of the NLS	1
	B. Purpose of this Study	2
II.	RESPONSE RATES, RESPONSE MODE, AND DATA QUALITY: DESCRIPTIVE STATISTICS	5
	A. Response Rates	5
	B. Mode of Response	9
	C. Quality of Response	14
III.	PREDICTING THIRD FOLLOW-UP RESPONSE	19
	A. General Approach	19
	B. Univariate Results	22
	C. Results of the Regression Analyses	25
	D. Discussion of Results	29
IV.	PREDICTING MAIL OR PERSONAL INTERVIEW RESPONSE TO THE THIRD FOLLOW-UP	30
	A. General Approach	30
	B. Results of the Discriminant Function Analyses	31
	C. Interpretation of Results	34
V.	PREDICTING THE QUALITY OF RESPONSE TO THE THIRD FOLLOW-UP	36
	A. General Approach	36
	B. Results	39
	C. Discussion of Results	42
VI.	THE EFFECTS OF THE INCENTIVE ON THE THIRD FOLLOW-UP	44
	A. Results of the Experimental Evaluation of the Incentive Check	44
	B. Mail Response Rate Implication for the Third Follow-Up Survey	46
	C. Cost Implications for the Third Follow-Up Survey	47
VII.	SUMMARY OF FINDINGS	50
	A. Third Follow-Up Response	50
	B. Mode of Response to the Third Follow-Up	50
	C. The Quality of Third Follow-Up Responses	51
	D. Importance of the Incentive	52
	References	53

CONTENTS--Continued

	Page
Appendixes	
A. Standardized Regression Results for the Prediction of Third Follow-Up Response	A.1
B. Summary of Discriminant Function Analyses for the Prediction of Mail or Personal Interview Response to the Third Follow-Up Survey	B.1
C. Regression Results for the Prediction of the Third Follow-Up Data Quality Indicators	C.1

Tables

1.	Response rates to the NLS instruments through the third follow-up survey	5
2.	Third follow-up response rates for subgroup defined by background variables	7
3.	Third follow-up response rates for subgroup defined by previous response patterns	8
4.	Mode of response to the follow-up surveys	9
5.	Mode of response to the third follow-up for various subpopulations defined by background variables	11
6.	Mode of response to the third follow-up for various subpopulations defined by the mode of response to previous surveys	13
7.	Percentage of third follow-up mail returns from respondents to first, second, and third follow-up surveys.	14
8.	Third follow-up data quality indices	15
9.	Average values of third follow-up data quality indices for various groups	16
10.	Average value of third follow-up quality indices for various subgroups defined by the mode of response to the follow-up surveys	18
11.	Variables used in predicting third follow-up response	20
12.	Means and intercorrelation of criterion and predictor variables	23
13.	Unstandardized regression results for the prediction of third follow-up response	26
14.	Standardized and unstandardized discriminant function coefficients	33
15.	Variables used in predicting third follow-up data quality indices not appearing elsewhere	37
16.	Signs of statistically significant regression coefficients associated with variables used to predict third follow-up data quality indices and final R^2 and F values associated with the regression equations	40
17.	Percentage of response over time	45
18.	Response rates to previous pilot studies and follow-up surveys	46

CONTENTS--Continued

Page

Appendix A Tables

Standardized regression results for the prediction of third follow-up response A.1

Appendix B Tables

Summary of discriminant function analyses for the prediction of the mode of response to the third follow-up survey B.1

Appendix C Tables

Regression results for the prediction of TFUORIND C.1
Regression results for the prediction of TFURIND C.2
Regression results for the prediction of TCSIND1 . . . . . C.3
Regression results for the prediction of TCSIND2 . . . . . C.4
Regression results for the prediction of TCIAIND C.5
Regression results for the prediction of TCIBIND C.6
Regression results for the prediction of TCICIND C.7
Regression results for the prediction of TCIDIND C.8
Regression results for the prediction of TCIEIND C.9
Regression results for the prediction of TCIFIND C.10

Figures

1. Plot of response mode group centroids on the two discriminant functions 31
2. Computation of cost savings in mail effects due to the incentive 48
3. Computation of cost savings in field interviews due to the incentive 49
4. Computation of total net savings due to the incentive 49

I. INTRODUCTION

This study was designed and implemented to examine the relationships among rate of response, data quality, different methods of survey data collection, and characteristics of potential respondents. It was anticipated that knowledge of these relationships would facilitate future data collection efforts. While the issues addressed and the answers suggested have some potential for generalization to a broad range of survey situations, the specific study purpose was focused on the National Longitudinal Study of the High School Class of 1972 (NLS). The specific vehicle for examining the issues was the NLS Third Follow-Up Survey.

This introductory section provides a brief overview of the scope of the NLS and the purpose of the study reported here. The remainder of the report is divided into six sections. Section II presents descriptive statistics concerning the incidence, mode, and quality of response to the Third Follow-Up Survey. Section III through V presents the results of modeling, by regression and discriminant function analyses, the incidence, mode, and quality of the third follow-up responses. Section VI contains a discussion of the effects of the incentive used in the Third Follow-Up Survey. All findings are summarized in Section VII.

A. General Description of the NLS

NLS is a large-scale survey sponsored by the National Center for Education Statistics (NCES), and its mission is to discover what happens to young people after they leave high school, as measured by their subsequent educational and vocational activities, plans, aspirations, and attitudes and to relate this information to their prior personal and educational experiences. Ultimately, the study is intended to allow a better understanding of the development of students as they pass through the American educational system and of the complex factors associated with individual educational and career outcomes. Such information is essential as a basis for effective planning, implementation, and evaluation of Federal policies and programs designed to enhance educational opportunity and achievement and to upgrade occupational attainments and career outcomes.

Following a rather extensive period of planning, which included the design and field test of survey instrumentation and procedures, a full-scale

survey was initiated in the spring of 1972 by the Educational Testing Service (ETS). The sample design called for a deeply stratified national probability sample of 1,200 schools with 18 seniors per school, school size permitting. The resulting base-year sample of 19,004 students from 1,061 high schools provided base-year data on up to three data collection forms--a Test Battery (TB), a Student Record Information Form (SRIF), and a Student Questionnaire (SQ). The key form, the SQ, was completed by 16,683 seniors.

Subsequent follow-up surveys have been carried out by Research Triangle Institute (RTI); three such follow-ups occurred between 1973 and 1977. The First Follow-Up Survey was conducted from October 1973 to April 1974. Added to the base-year sample were 4,450 seniors from the class of 1972 in 257 additional schools, some of which had been unable to participate earlier, bringing the total sample to 23,451 potential respondents. First follow-up forms were mailed to 22,654 students, and questionnaires were obtained for 21,350 of these. The Second Follow-Up Survey was conducted from October 1974 to April 1975. Questionnaires were mailed to 22,364 sample members, and completed questionnaires were obtained from 20,872. The Third Follow-Up Survey was conducted from October 1976 to May 1977. Questionnaires were mailed to 21,807 sample members whose addresses appeared sufficient and correct and who had not been removed from active status by prior refusal, reported death, or other reason. Some 20,092 completed questionnaires (about 92 percent) were returned.

In addition to the follow-up surveys carried out as an integral part of the NLS, a smaller sample of over 900 memebers of the high school graduating class of 1971 has been used for field testing instruments and operational procedures. The activities related to this pilot study sample have been known, collectively, as Concept 71.

B. Purpose of this Report

One of the significant features of the data collection processes employed by RTI during the follow-up surveys of the NLS sample has been the effort expended to maximize response rates. Such efforts have included extensive tracing activities, the mailing of questionnaires, newsletters, letters to parents and students, prompting postcards, and mailgrams. In the event that these approaches have failed to yield a mail response, sample members have been scheduled for field interviews. Telephone interviews have been used in

some instances for geographically isolated nonrespondents and to elicit key items of information when such items were incomplete or otherwise indeterminate on mail returns.

As might be expected from the variety of methods employed to collect data, consistently high response rates have been realized; however, considerable costs were involved. Thus, any information which might serve to identify subgroups of respondents for whom responses are more difficult, and therefore costly, to obtain could be valuable in planning subsequent surveys dealing with this specific population and could further facilitate the design of data collection procedures for similar longitudinal studies. Previous research has not provided fully adequate empirical bases for answering these questions; however, some inferences can be drawn from related studies. For example, it has been found that less educated people exhibit higher nonresponse rates in mail survey (Moser & Kalton, 1972) but are more likely to be locatable (Crider & Willits, 1973) and less likely to give contradictory responses (Lenski & Leggett, 1960). It has also been found that poor achievers in college exhibit a higher nonresponse rate (Neuss, 1943), and that nonresponse is greater among families of lower socioeconomic background (Vincent, 1964). None of these studies, however, have investigated the response patterns of major longitudinal surveys systematically .

Accordingly, the prediction of several outcome variables (including incidence of response to the NLS third follow-up, mode of response, and quality of response) is of interest. For each of these suggested outcomes, two classes of predictors may be utilized separately or in combination: (1) background variables such as race, sex, SES, ability, and high school program; and (2) variables from previous surveys analogous to the outcome under consideration. Information available from the first class alone could be of use in the design of other longitudinal survey's data collection procedures. Information based on both of these "predictor" classes may be used in developing procedures for any subsequent follow-up surveys of the NLS sample.

Another important set of issues for investigation focuses on the provision of a monetary incentive as part of the third follow-up data collection activities. Each subject was mailed a check for three dollars along with the first Third Follow-Up Questionnaire, under the assumption that an improvement in the mail response rate would result. The overall effect of this incentive did

appear to be positive, since the mail response rate increased from 69 percent for the second follow-up to 74 percent for the third follow-up; however, since all NLS sample members received the incentive, it is difficult to determine exactly how much of the increase was due to the use of the incentive for a number of reasons, including:

- a two-year time span between the second and third follow-ups (only one-year spans had separated the Base Year and First Follow-Up, and First and Second Follow-Up Surveys)
- a longer, somewhat more complex questionnaire that covered a period of two years rather than one and that included a few additional questions not previously asked of the sample members
- the fact that the response rate for the third follow-up should have been somewhat higher than the second follow-up as the base number of sample members had been reduced by 228 due to the elimination of refusals to the second follow-up and those sample members who had died or who, for physical or other reasons, were no longer eligible for participation.

To circumvent these problems in evaluating the effect of the incentives, use was made of the pilot study sample (Concept 71). The pilot study sample has been exposed to much the same kind and amount of mail data collection procedures, and of study information and materials as the NLS sample members. Consequently, the same ($3.00) incentive was offered to a stratified sample of one-half of the pilot study group during the fourth follow-up field test. The follow-up procedures used (except for the personal interview) were similar in nature and timing to those used for the NLS sample in the third follow-up. The full details of this incentive study have been reported elsewhere (King, 1979) and only a brief summary of the pertinent results are included in this report. References will be made to the incentive study as appropriate, to consider the potential effect of the incentive when interpreting patterns of response.

II. RESPONSE RATES, RESPONSE MODE, AND DATA QUALITY: DESCRIPTIVE STATISTICS

Three general topics are addressed in this section of the report. First, statistics are presented relating to response rates to the Third Follow-Up Survey for various groups defined by their background characteristics and their response pattern to prior surveys. The second topic considers the same fundamental data but focuses on whether the mode of response was by mail or through personal interview. The final topic concerns the quality of the responses which were obtained.

A. Response Rates

Table 1 presents the overall response rates for the base year, first, second, and third follow-up instruments based on the fully augmented sample size of 23,451 subjects. Considering the span of time encompassed by the study and the growing diversity of the activities in which the members of the

Table 1.--Response rates to the NLS instruments through the third follow-up survey

Administration	Response rates
Base year	.711[a/]
First follow-up	.910
Second follow-up	.890
Third follow-up	.857

[a/] The very low Base Year response rate was due primarily to the refusal of sampled schools to participate in the collection of data from students. School refusal was associated with the late start of field data collection.

NOTE.--Based on the availability of completed instruments for the full sample of 23,451 individuals. These statistics include as part of the population those who refused to participate, were deceased, institutionalized or otherwise unable to respond. Thus, these are the most conservative rates which could have been computed.

high school class of 1972 were involved, the success of the NLS follow-up efforts has been quite good. With the exception of the relatively low base year response rate, which was due primarily to the refusal of schools to permit the administration of the student questionnaire, response rates have remained above 85 percent. At the same time, however, there has been a discernable trend toward increased nonresponse as would be expected in any panel study taking place over so long a period.

The third follow-up response rates for subgroups defined by the background variables of aptitude, SES, ethnicity, sex, and high school curriculum are provided in Table 2. With the exception of the categories labeled "unknown," the response rates were quite high (84 percent or better) for all of the subgroups, suggesting only minor potential bias, with respect to the background variables considered, among the responding group. The response rates for those in the "unknown" groups varies considerably depending on the particular classification variable involved but is lower than that for any other subgroup in a given classification. Much of this reflects the fact that consistent nonresponse to the base year and follow-up surveys usually was necessary for a case to be "unclassified." Thus, the difference reflects response history much more than classification by background variable (see subsequent discussion). The only exception involves aptitude, which was measured only in the base year and which, therefore, could not have been made available through later responses to the follow-up surveys (particularly relevant here are those added to the sample subsequent to the base year).

Table 3 provides response rates for subgroups defined by prior response history, and these values show much more dramatic subgroup difference. For those sample members who had responded to all three prior instruments, the response rate approached 95 percent. At the other extreme a response rate of less than 5 percent was obtained among those who had responded previously to only the Base Year Questionnaire. The low response rates shown in Table 3 are likely due, to some extent, to the inclusion in the statistics of those who were defined to be ineligible for the third follow-up survey through loss of contact, death, other incapacitation or personal request to be removed from the study. (In fact, Third Follow-Up Questionnaires were not mailed to such individuals.) Moreover, the rates are confounded somewhat by a methodology that added sample members after the base year, who obviously had no opportunity for base year response. Beyond effects that may be attributable to

Table 2.--Third follow-up response rates for subgroup defined by background variables

Classification	Subgroup	Number of cases	Third follow-up response rate
Total		23,451	.857
Aptitude	Low	4,798	.844
	Middle	7,008	.898
	High	4,054	.933
	Unknown	7,591	.786
SES	Low	6,451	.871
	Middle	10,654	.894
	High	5,087	.911
	Unknown	1,259	.249
Ethnicity	Black	3,119	.860
	White	17,159	.902
	Hispanic	986	.844
	Other	1,100	.858
	Unknown	1,087	.157
Sex	Male	11,167	.875
	Female	11,336	.899
	Unknown	948	.145
High school curriculum	General	8,172	.872
	Academic	8,937	.909
	Voc-Tech	5,481	.877
	Unknown	861	.041

NOTE.--Based on the availability of completed instruments for the full sample of 23,451 individuals. There statistics include as part of the population those who refused to participate, were deceased, institutionalized or otherwise unable to respond. Thus, these are the most conservative rates which could have been computed.

such procedural artifacts the results shown in Table 3 still suggest that third follow-up response may be attributed to individual response proclivities as reflected in prior patterns of respondings. Two related facets of prior response pattern seem to be involved: (1) consistency of response (i.e, the number of prior responses); and (2) recency of response.

Table 3.--Third follow-up response rates for subgroup defined by previous response patterns

Response pattern for prior questionnaires			Number of cases	Third follow-up response rate
Base year	First follow-up	Second follow-up		
No	No	No	919	.131
No	No	Yes	134	.881
No	Yes	No	421	.266
No	Yes	Yes	5,294	.926
Yes	No	No	504	.036
Yes	No	Yes	544	.879
Yes	Yes	No	735	.314
Yes	Yes	Yes	14,900	.947

NOTE.--Based on the availability of completed instruments for the full sample of 23,451 individuals. These statistics include as part of the population those who refused to participate, were deceased, institutionalized for otherwise unable to respond. Thus, these are the most conservative rates which could have been computed.

The effect of response consistency can be observed by aggregating the response rates given in Table 3 for those responding to none, one, two, or three of the prior questionnaires. Third follow-up response rates for such groups are .131, .234, .854, and .947, respectively. The importance of the recency of response as a predictor of third follow-up response also can be seen through an aggregation of Table 3 response pattern statistics into subgroups defined by response to the base year, first and second follow-up surveys separately. The strength of the association between prior response and third follow-up response increases with the recency of the data collection effort. There is only a difference of .11 in response rate between those who did (.89)

and did not (.78) respond to the Base Year Questionnaire; a difference of .56 is obtained between those who did (.91) and did not (.35) respond to the first follow-up, and, finally a difference of .75 exists between the subgroups defined by response to the second follow-up (.94 for respondents, .19 for nonrespondents).

While response recency and consistency are somewhat confounded in the results reported above, it is possible to observe the recency effect controlling for consistency within the groups that responded to only one or to two of the prior surveys. The three groups responding to only one of the prior surveys are the second-, third-, and fifth-listed groups in Table 3. Ordering these groups by the recency of the one prior response, the third follow-up response rates are .88, .27, and .04, respectively. The fourth-, sixth-, and seventh-listed groups in Table 3 responded to two prior surveys. Ordering these groups by recency of response (inversely by the single prior nonresponse), the third follow-up response rates are .93, .88, and .31, respectively.

B. Mode of Response

Not only has the response rate for the various follow-up surveys remained quite high, but responses have been obtained increasingly through the mail data collection effort. As can be seen in Table 4, over 2,000 more mail responses were received for the third follow-up than for the first. Con-

Table 4.--Mode of response to the follow-up surveys

Questionnaire	Mode of response					
	Mail		Personal interview		Nonresponse	
First follow-up	14,019	(59.8)	7,331	(31.3)	2,101	(9.0)
Second follow-up	15,058	(64.2)	5,814	(24.8)	2,579	(11.0)
Third follow-up	16,102	(68.7)	3,990	(17.0)	3,359	(14.3)

NOTE.--All entries consist of numbers of individuals and percentages (in parentheses), which sum to 100 for a given row, within rounding error.

currently, the number of personal interviews declined by about half, from 7,331 to 3,990. This, coupled with the slow increase in nonresponse, has led to a considerable change in the make-up of the respondent pool vis a vis how the data were collected. While the mail responses on the first follow-up accounted for just under two-thirds of the total received, they accounted for over four-fifths of the data on the third follow-up. Given the much greater costs for personal interview data collection, this shift clearly marks the third follow-up as far more cost effective than the first.*

Statistics on the mode of response to the third follow-up for various subpopulations defined by the same background variables considered previously are presented in Table 5. This table reveals some trends of potential interest to those responsible for surveys of a similar nature. Having considered trends related to response/nonresponse in the previous section, only those aspects having to do with mail or personal interview response will be considered here. Although comparisons of mail versus interview on the basis of percentages given in Table 5 may be somewhat misleading due to differential nonresponse rates (particularly so for the "unknown" groups), the basic trends suggested by the unconditional percentages are also generally supported by percentages that are conditional on the responding subset.†

From Table 5 it can be observed that there is a positive relationship between response by mail and aptitude (which is probably reflected, to some extent, in the high school program classification). Those in the high aptitude group were more likely to respond to the third follow-up by mail than were those in the middle or low groups. Somewhat similar but less pronounced results were obtained for the subgroups defined by SES; those in the high SES group show a higher likelihood of responding by mail than do those in the middle or low groups. Responses from white sample members were more likely to be obtained through the mail survey effort than were those from members of other ethnic groups. Subgroups defined by high school program and sex also reflect moderate differences along these lines.

---

* As shown in the final section of the report, much of this shift must be credited to the use of the incentive check on the third follow-up.

† Such conditional percentages are directly computable from the values given in Table 5.

Table 5.--Mode of response to the third follow-up for various subpopulations defined by background variables

Classification	Subgroup	Mail	Personal interview	Nonresponse
Total		16,102 (68.9)	3,990 (17.0)	3,359 (14.3)
Aptitude	Low	2,947 (61.4)	1,104 (23.0)	747 (15.6)
	Middle	5,246 (74.9)	1,047 (14.2)	715 (10.2)
	High	3,334 (82.2)	449 (11.1)	271 (6.7)
	Unknown	4,575 (60.3)	1,390 (18.3)	1,626 (21.4)
SES	Low	4,248 (65.9)	1,370 (21.2)	833 (12.9)
	Middle	7,694 (72.2)	1,830 (17.2)	1,130 (10.6)
	High	3,903 (76.7)	733 (14.4)	451 (8.9)
	Unknown	257 (20.4)	57 (4.5)	945 (75.1)
High school program	General	5,564 (68.1)	1,565 (19.2)	1,043 (12.8)
	Academic	6,749 (75.5)	1,373 (15.4)	815 (9.1)
	Voc-Tech	3,761 (68.6)	1,045 (19.1)	675 (12.3)
	Unknown	28 (3.2)	7 (0.8)	826 (95.9)
Ethnicity	Black	1,705 (54.7)	977 (31.3)	437 (14.0)
	White	12,895 (75.2)	2,568 (15.0)	1,696 (9.8)
	Hispanic	612 (62.1)	220 (22.3)	154 (15.6)
	Other	739 (67.2)	205 (18.6)	156 (14.2)
	Unknown	151 (13.9)	20 (1.8)	916 (84.3)
Sex	Male	7,538 (67.5)	2,228 (20.6)	1,401 (12.5)
	Female	8,430 (74.4)	1,759 (15.5)	1,147 (10.1)
	Unknown	134 (14.1)	3 (.3)	811 (85.5)

NOTE.--All entries consist of numbers of individuals and percentages (in parentheses), which sum to 100 for a given row, within rounding error.

The importance of the extensive personal interview effort in reducing response biases among these various subgroups is also clear from these statistics. In the absence of this effort responses would have been far more disproportionately available from the high aptitude, high SES and white subgroups of the NLS sample than was, in fact, the case. Instead of responses from white sample members being almost half again as likely as those from black sample members (which was the case for the mail portion of the follow-up effort), they were only seven percent more likely. Similar but smaller effects can be seen with respect to the other major background variables.

Table 6 displays the mode of response rates for the subgroups defined by their mode of response (or lack thereof) to the Base Year, First and Second Follow-Up Surveys. Although some response mode patterns are evident in Table 6, the inclusion of the nonresponding groups (and differential nonresponse within groups)* tends to mask the patterns somewhat. For that reason, mode of response data are summarized in Table 7, in which mail return percentages conditional to only to respondents to first, second, and third follow-up studies are considered. (Due to the nature of the base year sampling and mode of data collection--e.g., not including the full augmented sample, and collection of questionnaire data through in-school group administration--base year mode of response is not considered and results are partitioned by base year response.) From Table 7, the strong effects of consistency (number of prior responses by mail) and recency (latest follow-up returned by mail) on the likelihood of response by mail to the third follow-up are clearly observable. On the other hand, response to the base year study appears to be unrelated to the likelihood of responding by mail to the third follow-up.

* This has implications both from the standpoint of individual motivation to participate in the study and the availability of current address information to which a mailed questionnaire could be sent. Additional comments regarding the effect of this classification scheme is presented in Section IV.

Table 6.--Mode of response to the third follow-up for various subpopulations defined by the mode of response to previous surveys

Base year	First follow-up	Second follow-up	Mode of response to the third follow-up: Mail response	Personal interview	Nonresponse
Non-Response	Mail	Mail	2,082 (86.2)	274 (11.3)	59 (2.4)
		Personal interview	388 (62.0)	183 (29.2)	55 (8.8)
		Nonresponse	22 (19.6)	21 (18.7)	69 (61.6)
	Personal interview	Mail	805 (76.4)	181 (17.2)	68 (6.5)
		Personal interview	481 (40.1)	509 (42.5)	209 (17.4)
		Nonresponse	27 (8.7)	42 (13.6)	240 (77.7)
	Nonresponse	Mail	48 (69.6)	14 (20.3)	7 (10.1)
		Personal interview	41 (63.1)	15 (23.1)	9 (13.8)
		Nonresponse	120 (13.1)	0 (0.0)	799 (86.9)
Response	Mail	Mail	8,172 (89.1)	796 (8.7)	206 (2.2)
		Personal interview	921 (64.0)	392 (27.2)	127 (8.8)
		Nonresponse	70 (27.8)	32 (12.7)	150 (59.5)
	Personal interview	Mail	1,652 (74.3)	428 (19.3)	142 (6.4)
		Personal interview	836 (40.5)	915 (44.3)	313 (15.2)
		Nonresponse	64 (13.3)	65 (13.5)	354 (73.3)
	Nonresponse	Mail	79 (63.7)	26 (21.0)	19 (15.3)
		Personal interview	276 (65.7)	97 (23.1)	47 (11.2)
		Nonresponse	18 (3.6)	0 (0.0)	486 (96.4)

NOTE.--All entries consist of numbers of individuals and percentages (in parentheses), which sum to 100 for a given row, within rounding error.

Table 7.--Percentage of third follow-up mail returns from respondents to first, second, and third follow-up surveys

First follow-up mode of response	Second follow-up mode of response	Base year response	
		Responded	Did not respond
Mail	Mail	91.1% (N=8,968)	88.4% (N=2,356)
Mail	Interview	70.1 (N=1,313)	68.0 (N=571)
Interview	Mail	79.4 (N=2,080)	81.6 (N=986)
Interview	Interview	47.7 (N=1,751)	48.6 (N=990)

NOTE.--Cell entries represent the percentage of third follow-up respondents in that cell who responded by mail; the number of cases on which the percentage is based is also provided in parentheses.

C. Quality of Response

The quality of data obtained from the third follow-up respondents is also of concern. Since a response to the questionnaire was required for the definition of data quality, all statistics presented are based on the 20,092 sample members who responded to the third follow-up. The ten data quality indices available for the third follow-up responses are presented in Table 8 together with a brief description of their definition.* Table 9 displays the overall means for these indices as well as those for the subgroups defined by the third follow-up mode of response and other background variables. The first four of these indices tend to have lower values for the higher quality responses; conversely, the remaining six higher quality responses are indicated by higher values. Keeping this in mind, it is clear from an inspection of the table that the more expensive personal interview effort yields better information

* More detailed information on these indices is available in the Data File Users Manual (Levinsohn, et al, 1978).

Table 8.--Third follow-up data quality indices

Variable name	Explanation
TFURIND	Third follow-up routing index (low values imply few routing errors)
TFUORIND	Third follow-up out-of-range index (low values imply few out-of-range errors)
TCSIND1	Third follow-up key question consistency index (low values imply few key questions inconsistencies)
TCSIND2	Third follow-up record quality consistency index (low values imply few record quality inconsistencies)
TCIAIND	Third follow-up Section A completion index [general information] (expressed as a percentage of eligible questions answered)
TCIBIND	Third follow-up Section B completion index [work experience] (expressed as a percentage of eligible questions answered)
TCICIND	Third follow-up Section C completion index [education and training] (expressed as a percentage of eligible questions answered)
TCIDIND	Third follow-up Section D completion index [military service] (expressed as a percentage of eligible questions answered)
TCIEIND	Third follow-up Section E completion index [family status] (expressed as a percentage of eligible questions answered)
TCIFIND	Third follow-up Section F completion index [experience and opinions] (expressed as a percentage of eligible questions answered)

Table 9.--Average values of third follow-up data quality indices for various groups

Classification	Subgroups	Number of cases	TFURIND*	TFUORIND*	TCSIND1*	TCSIND2*	TCIAIND†	TCIBIND†	TCICIND†	TCIDIND†	TCIEIND†	TCIFIND†
Total		20,092	1.92	.023	.701	.884	95.28	90.28	81.70	97.96	86.10	88.47
Mode of response	Mail	16,102	2.13	.027	.800	.923	96.54	90.50	80.95	97.92	85.21	89.20
	Personal interview	3,990	1.10	.006	.303	.728	90.20	89.39	84.74	98.13	89.72	85.54
Aptitude	Low	4,051	2.80	.034	.854	1.151	93.14	86.28	79.64	96.64	83.56	86.38
	Middle	6,293	1.61	.020	.611	.786	96.22	91.90	82.84	98.63	86.83	89.42
	High	3,783	1.24	.018	.630	.665	96.94	93.32	83.25	98.86	88.72	89.76
SES	Low	5,618	2.41	.028	.882	1.031	94.26	88.31	80.94	96.93	85.06	87.59
	Middle	9,524	1.82	.020	.643	.835	95.59	90.89	82.12	98.28	86.34	88.84
	High	4,636	1.44	.021	.650	.765	96.16	91.83	82.04	98.81	87.19	89.07
High school program	General	7,129	2.13	.024	.710	.936	94.66	89.27	81.70	97.64	85.54	88.08
	Academic	8,122	1.52	.021	.704	.803	96.24	91.79	81.88	98.60	87.51	89.14
	Voc-Tech	4,806	2.30	.026	.683	.941	94.56	89.26	81.42	97.42	84.60	87.89
Ethnicity	Black	2,682	3.09	.030	1.088	1.253	92.72	85.46	76.94	95.86	84.45	86.39
	White	15,463	1.65	.020	.613	.791	95.86	91.37	82.80	98.49	86.49	89.02
	Hispanic	832	2.48	.019	.889	1.198	93.56	87.89	79.34	95.96	85.46	86.91
	Other	944	2.24	.041	.753	.953	94.78	89.21	80.74	98.06	86.24	87.39
Sex	Male	9,766	1.92	.028	.766	.857	95.01	90.44	81.05	97.24	86.64	87.84
	Female	10,189	1.89	.017	.629	.900	95.56	90.21	82.44	98.73	85.69	89.14

* Low values indicate high quality.

† High values indicate high quality.

as measured by the first four indicators of data quality (involving routing, out-of-range, and consistency errors). On the other hand, the six completeness indices display more of a mixed result with half being higher for the personal interview data and half lower. The completion index differences could be artifactual as a result of administering only selected "key items" of the questionnaire when the interview was conducted by telephone. This may also have arisen from the fact that those who responded by personal interview were less willing to respond to some items in a face-to-face situation than were those who responded by mail. While the interviewers could avoid the response inconsistencies measured by the first four indices, they could not necessarily guarantee that the subjects would provide complete information to all of the questions in the various sections of the questionnaire.

Considering subgroups defined by background variables, it is clear that meaningful group differences in data quality do exist. There is a clear positive relationship between aptitude and data quality; those in the high aptitude group show the highest data quality while those in the lowest aptitude group show the least, for all but one of the indicators considered. Similar relationships can be observed for subgroups defined by SES, high school program, and ethnicity. Among background variables, data quality varies least between subgroups defined by sex.

Table 10 contains data quality statistics for subgroups defined by pattern of response to the follow-up surveys. In general, the data quality as measured by the first four indicators was better for the groups which provided interview responses to the third follow-up, with relatively little variability among these groups. Among those responding by mail, however, more of a pattern in the variability of the data quality can be seen. The nature of this pattern suggests that those with the greatest amount of experience in responding to the follow-up surveys by mail tended to provide the highest quality data on the third follow-up. At the opposite extreme, those with very little experience in responding to the follow-up surveys by themselves (typified by those who had responded to neither previous follow-up) provided the poorest quality data.

Table 10.--Average value of third follow-up quality indices for various subgroups defined by the mode of response to the follow-up surveys

Mode of response to follow-ups			Quality indices									
First	Second	Third	N	TFURIND*	TFUORIND*	TCSIND1*	TCSIND2*	TCIAIND†	TCIBIND†	TCICIND†	TCIDIND†	TCIEIND†	TCIFIND†
Mail	Mail	Mail	10,254	1.74	.021	.701	.803	97.19	92.29	82.20	98.59	86.56	90.04
Mail	Mail	Interview	1,070	.99	.008	.327	.733	90.23	89.68	84.38	98.62	90.45	85.67
Mail	Interview	Mail	1,309	2.64	.031	.981	1.074	96.25	88.44	79.20	96.74	82.72	87.77
Mail	Interview	Interview	575	.98	.007	.283	.757	89.36	88.54	83.53	98.15	89.07	84.76
Mail	NR	Mail	92	3.53	.076	1.620	1.402	96.44	87.51	78.95	93.55	83.48	87.46
Mail	NR	Interview	53	1.72	.000	.491	.622	91.49	91.49	85.00	93.26	89.94	87.53
Interview	Mail	Mail	2,457	2.39	.035	.785	1.007	95.78	88.39	80.66	97.85	83.98	88.48
Interview	Mail	Interview	609	1.32	.003	.343	.767	90.48	89.36	84.48	97.68	89.96	85.99
Interview	Interview	Mail	1,317	3.46	.035	1.084	1.304	93.93	84.59	76.10	96.04	81.54	86.53
Interview	Interview	Interview	1,424	1.10	.005	.268	.665	90.75	89.77	85.86	98.61	89.47	85.80
Interview	NR	Mail	91	4.09	.121	2.418	2.253	92.22	82.00	65.79	88.74	70.51	81.77
Interview	NR	Interview	107	1.10	.000	.243	1.122	85.95	86.64	82.50	96.36	88.11	82.22
NR	Mail	Mail	127	2.78	.031	1.063	1.087	95.47	89.80	77.78	94.07	84.24	89.35
NR	Mail	Interview	40	1.70	.000	.350	1.050	85.18	85.50	78.02	90.42	88.02	80.28
NR	Interview	Mail	317	2.62	.022	1.060	1.032	96.27	88.38	79.99	96.96	84.38	89.36
NR	Interview	Interview	112	1.39	.009	.366	.696	91.15	89.19	85.83	96.29	89.87	86.44
NR	NR	Mail	138	4.44	.167	1.572	1.645	93.51	84.88	72.81	93.86	78.91	83.20
NR	NR	Interview	0	-	-	-	-	-	-	-	-	-	-
Total			20,092	1.92	.023	.701	.884	95.28	90.28	81.70	97.96	86.10	88.47

* Low values indicate high quality.

† High values indicate high quality.

III. PREDICTING THIRD FOLLOW-UP RESPONSE

There were two necessary conditions which must have been met for the survey staff to have obtained a response from any given sample member. The first is that the sample member must have been contacted by the survey staff, either through the mail data collection effort or, failing that, through the personal interview process. The second is that, given receipt of the questionnaire, the sample member must have responded to and returned the survey instrument. The net result of these requirements is that response to the third follow-up survey was dependent upon a mixture of personal and procedural factors. The constructs thought to be the best indicators of these personal and procedural factors were the sample members' background and prior follow-up response history. Based on these considerations four basic questions, listed below, motivated the analyses presented in this section of the report.

(1) How well can third follow-up response be predicted from knowledge of response to the second follow-up alone?

(2) To what extent does knowledge of response to earlier surveys improve this prediction?

(3) Can this prediction be improved upon through the addition of interaction terms involving prior response and type of response?

(4) Does knowledge of background variables (such as SES, ethnicity, sex, and high school program) permit further improvement in prediction over that already obtained?

A. General Approach

These questions were addressed through a series of regression analyses using a binary variable representing third follow-up response as the criterion. The variables used as predictors are presented in Table 11, with brief explanations of their meaning, specifications for the assignment of values and mnemonic labels used in subsequent discussion. The first analysis considered only the indicator of second follow-up nonresponse (SFUNR) as a predictor. For the second analysis, four additional predictors were added to the regression equation: the indicators of base year response (BYR), first follow-up response through personal interview (FFUPI), first follow-up nonresponse (FFUNR), and second follow-up response through personal interview (SFUPI). For the third analysis, the regression equation was further augmented by the addition of the

Table 11.--Variables used in predicting third follow-up response

Variable name	Explanation	Definition of values
BYR	Indicator of base year response	0 = Did not respond to base year questionnaire 1 = Responded to base year questionnaire
FFUPI	Indicator of first follow-up response through personal interview	0 = Did not respond to first follow-up questionnaire through personal interview 1 = Responded to first follow-up questionnaire through personal interview
FFUNR	Indicator of first follow-up nonresponse	0 = Responded to first follow-up 1 = Did not respond to first follow-up
SFUPI	Indicator of second follow-up response through personal interview	0 = Did not respond to second follow-up through personal interview 1 = Responded to second follow-up through personal interview
SFUNR	Indicator of second follow-up nonresponse	0 = Responded to second follow-up 1 = Did not respond to second follow-up
FFUPI*SFUPI	Interaction term for personal interview response to both the first and second follow-ups	0 = Did not respond to both follow-ups by personal interview 1 = Responded to both follow-ups by personal interview
FFUPI*SFUNR	Interaction term for personal interview response to the first follow-up and nonresponse to the second	0 = Did not respond by personal interview to the first follow-up and not respond to the second 1 = Responded by personal interview to the first follow-up and did not respond to the second

(continued)

Table 11.--Variables used in predicting third follow-up response--Continued

Variable name	Explanation	Definition of values
FFUNR*SFUPI	Interaction term for nonresponse to the first follow-up and personal interview response to the second	0 = Did not fail to respond to the first follow-up and respond by personal interview to the second 1 = Did not respond to the first follow-up and responded by personal interview to the second
FFUNR*SFUNR	Interaction term for nonresponse to both first and second follow-ups	0 = Did not fail to respond to both the first and second follow-ups 1 = Did not respond to either the first or the second follow-up
SEX	Sex indicator	0 = Male 1 = Female
SESM	Indicator of medium level of SES	0 = Not medium level of SES 1 = Medium level of SES
SESH	Indicator of high level of SES	0 = Not high level of SES 1 = High level of SES
HPMGEN	Indicator of general high school program	0 = Not general high school program 1 = General high school program
HPMACAD	Indicator of academic high school program	0 = Not academic high school program 1 = Academic high school program
ETHBLK	Indicator of black ethnicity	0 = Not black ethnicity 1 = Black ethnicity
ETHWHT	Indicator of white ethnicity	0 = Not white ethnicity 1 = White ethnicity

NOTE.--The Base Year Questionnaire was administered only in the schools attended by the subjects at that time; therefore, only one mode of data collection was used. In addition, Base Year nonresponse was due almost exclusively to schools refusing to cooperate. Since this study focused primarily on individual characteristics and follow-up practices as prediction of response, no interaction terms involving the Base Year response were used.

four first-order interaction terms associated with the pattern of follow-up response (FFUPI*SFUPI, FFUPI*SFUNR, FFUNR*SFUPI, and FFUNR*SFUNR). The final analysis added the background variables to the regression equation.

Despite previous efforts to maximize the completeness of the data over the various collection efforts some background information was missing, principally for sample members who had never responded to any of the questionnaires.* Instead of using strategies involving the imputation of missing data or weight adjustment techniques, cases with missing data were simply eliminated for these analyses. This had the effect of reducing the number of cases to 22,067 (as compared to the 23,451 cases on which the descriptive statistics are based). From the standpoint of the original population this obviously introduces some biases in the regression statistics (of an undetermined magnitude and direction); however, the number of cases remaining is sufficiently large for this group to constitute a legitimate population of interest. As such, the reader must keep in mind the inherent limitations on the extent to which the results presented here may be generalized.

B. Univariate Results

Means and intercorrelations of all variables used in this set of analyses are presented in Table 12. Since all of the variables were dichotomies, inspection of the mean values effectively reveals the proportion of the cases having the characteristic assigned unit value. For instance, the mean value of the indicator of base year response, BYR, was .75 indicating that, for the 22,067 cases used in this analysis, 75 percent responded to the Base Year Questionnaire. Likewise, the mean value of .02 for the interaction term indicating nonresponse to both the First and Second Follow-Up surveys, FFUNR*SFUNR, means that only two percent of these cases failed to respond to both of the first two follow-ups.

As can be determined by comparing the mean values of the background variables with the distributions implicit in Table 2, the characteristics of these subjects were nearly identical to those from whom each individual variable was available. (Given the procedure of exclusing "unclassifiable" cases,

* Aptitude information was available only for those who had responded to the Base Year Questionnaire. Rather than attempting to impute missing values, the variable was simply excluded from these analyses.

Table 12.--Means and intercorrelation of criterion and predictor variables

	Mean	TFUR	BYR	FFUPI	FFUNR	SFUPI	SFUNR	FFUPI*SFUPI	FFUPI*SFUNR	FFUNR*SFUPI
TFUR	0.89	1.00000								
BYR	0.75	-0.00501	1.00000							
FFUPI	0.32	-0.13999	-0.13993	1.00000						
FFUNR	0.04	-0.28299	0.12464	-0.14945	1.00000					
SFUPI	0.26	-0.04791	-0.08940	0.30813	0.07851	1.00000				
SFUNR	0.06	-0.56868	0.03731	0.07411	0.34828	-0.15428	1.00000			
FFUPI*SFUPI	0.15	-0.07330	-0.10934	0.60021	-0.08970	0.70183	-0.10828	1.00000		
FFUPI*SFUNR	0.03	-0.35248	-0.01990	0.25050	-0.03744	-0.10230	0.66307	-0.07180	1.00000	
FFUNR*SFUPI	0.02	-0.00283	0.07949	-0.09532	0.63779	0.23340	-0.03601	-0.05721	-0.02388	1.00000
FFUNR*SFUNR	0.02	-0.40421	0.08325	-0.09982	0.66789	-0.08536	0.55329	-0.05991	-0.02500	-0.01992
SESH	0.48	0.01226	-0.00955	-0.00927	0.00411	-0.01265	-0.00146	-0.01606	0.00283	0.00570
SESH	0.23	0.03529	0.01311	-0.10361	-0.02779	-0.07770	-0.02342	-0.08383	-0.01881	-0.01079
HPHGEN	0.36	-0.03746	-0.07473	0.09040	0.02264	0.05556	0.03917	0.05770	0.02839	0.01079
HPHACAD	0.40	0.05382	0.04080	-0.13299	-0.02870	-0.08840	-0.05980	-0.09206	-0.04422	-0.00877
ETHBLK	0.14	-0.04104	-0.05176	0.13406	0.01350	0.15143	0.02534	0.14862	0.03176	0.01385
ETHMIT	0.77	0.06237	0.01949	-0.14001	-0.03086	-0.14409	-0.04085	-0.14342	-0.03118	-0.01577
SEX	0.51	0.03663	-0.00319	-0.08985	-0.04007	-0.07262	-0.03287	-0.08632	-0.02884	-0.02292

(continued)

Table 12.--Means and intercorrelations of criterion and predictor variables--Continued

	FFUNR*SFUNR	SESH	SESH	HPHGEN	HPHACAD	ETHBLK	ETHWHT	SEX
SFUNR*SFUNR	1.00000							
SESH	0.00501	1.00000						
SESH	-0.02018	-0.52365	1.00000					
HPHGEN	0.01295	0.01204	-0.11235	1.00000				
HPHACAD	-0.02639	-0.02037	0.28087	-0.60998	1.00000			
ETHBLK	0.00362	-0.13498	-0.16775	0.01218	0.07425	1.00000		
ETHWHT	-0.02800	0.15128	0.20418	-0.03694	0.10749	-0.72930	1.00000	
SEX	-0.01021	0.00080	-0.04449	-0.03126	-0.04695	0.04473	-0.02955	1.00000

the relatively low rate of unclassifiables, and the large extent of overlap among unclassifiables for each variable considered, such similarity would certainly be expected.) On the other hand, when the mean values for the questionnaire response (TFUR) and nonresponse variables (FFUNR and SFUNR) are compared with the information presented in Table 1, it is clear that the subjects whose data were used in these analyses were about 5 percentage points more likely to have responded to each of the questionnaires. (Given the much lower response rates among the excluded unclassifiables, this is to be expected also.)

The zero-order intercorrelations among all of the variables are also provided in Table 12. As an example of how these statistics should be interpreted, the correlation between the indicator of response to the third follow-up, TFUR, and the indicator of nonresponse to the second follow-up, SFUNR, was approximately -.57. This means that nonresponse to the second follow-up indicates a lowered likelihood of response to the third. Similarly, the correlation between the indicator of nonresponse to the first follow-up, FFUNR, and the indicator of nonresponse to the second, SFUNR, of approximately .31 means that those who did not respond to the first were also more likely not to respond to the second and vice versa. These intercorrelations together with the means and standard deviations, basically define the results of the regression analyses discussed below.

There are several points concerning the correlations involving TFUR that should be noted. The first is that the single largest correlation, in an absolute sense, is with the indicator of second follow-up nonresponse, SFUNR. This lends support to the idea that the most recent information would be of the greatest use in predicting third follow-up response. Coincidentally, the variable with which TFUR had one of the lowest correlations was the indicator of base year response, BYR. Apart from this, it was generally the case that the variables associated with questionnaire response were more highly related to third follow-up response than were the variables related to background.

C. Results of the Regression Analyses

The results from all four of the regression analyses are summarized in Table 13. This table contains the unstandardized regression coefficients for each model, the F statistics associated with each estimated coefficient, the R^2 value for each regression equation, and their associated F statistics. The table also contains the extent to which each successive regression equation

Table 13.--Unstandardized regression results for the prediction of third follow-up response

Variable	Equation 1		Equation 2		Equation 3		Equation 4	
	B	F	B	F	B	F	B	F
BYR	-	-	.005	1.73	.005	1.45	.005	1.75
FFUPI	-	-	-.055	197.04[a]	-.041	70.53[a]	-.040	66.19[a]
FFUNR	-	-	-.148	269.88[a]	-.133	35.59[a]	-.131	34.38[a]
SFUPI	-	-	-.068	268.58[a]	-.065	121.11[a]	-.064	118.22[a]
SFUNR	-.749	10546.59[a]	-.682	8116.78[a]	-.568	1642.40[a]	-.567	1631.63[a]
FFUPI*SFUPI	-	-	-	-	-.030	12.97[a]	-.031	13.05[a]
FFUPI*SFUNR	-	-	-	-	-.099	31.57[a]	-.100	32.33[a]
FFUNR*SFUPI	-	-	-	-	.107	17.25[a]	.106	16.63[a]
FFUNR*SFUNR	-	-	-	-	-.242	71.77[a]	-.244	72.32[a]
SEX	-	-	-	-	-	-	.001	.05
SESH	-	-	-	-	-	-	.010	5.60[a]
SESH	-	-	-	-	-	-	.008	2.85
HPHGEN	-	-	-	-	-	-	-.003	.46
HPHACAD	-	-	-	-	-	-	-.004	.94
ETHBLK	-	-	-	-	-	-	.017	5.53[a]
ETHWHT	-	-	-	-	-	-	.014	5.84[a]
Constant	.935	-	.992	-	.988	-	.971	-
R^2(F)	.32340	(10546.59)[a]	.35338	(2411.23)[a]	.36203	(1390.76)[a]	.36245	(783.48)[a]
R^2 change (F)[b]	.32340	(10546.59)[a]	.02998	(188.94)[a]	.00865	(74.77)[a]	.00042	(2.08)

[a] Statistically significant with $\alpha = .01$.

[b] The F value for the R^2 change associated with Equation 1 is simply that of the R^2 value of the model. The others are for the changes in R^2 attributable to the additional variables incorported and were computed using the following formula:

$$F = \frac{(R_g^2 - R_s^2)/(k_g - k_s)}{(1 - R_g^2)/(N-k_g-1)}$$

Where R_g^2 refers to the larger of the two R^2 values used to compute the change in R^2, k_g is the number of predictor variables in the equation with the larger R_g^2 and N is the number of cases used in the analysis (22,067).

increases the R^2 value over the one preceeding it and an F statistic associated with each R^2 change. The standardized regression coefficients are contained in Appendix A.

The solution to the first equation, in which third follow-up response was predicted from second follow-up nonresponse alone, provides a benchmark against which the rest of the regression analyses may be compared. The regression coefficient associated with SFUNR differs statistically from zero, and overall, the regression equation accounts for just over 32 percent of the variance in the criterion variable (i.e., $R^2 = .32$). Substantively, the value of the constant in the equation indicates that the predicted response rate among those for whom the value of SFUNR was zero (i.e., those who did respond to the second follow-up) is 93.5 percent. On the other hand, the predicted response rate for those who had a value of SFUNR of one (i.e., those not responding to the second follow-up) is 18.6 percent (i.e., 74.9 percent lower).

The second regression equation includes four additional variables in the model (the indicators of base year response, first follow-up nonresponse and personal interview response, and second follow-up personal interview response); for this model the proportion of third follow-up response variance accounted for is statistically larger even though the increase is not large (about .03) in an absolute sense. With the exception of the coefficient associated with the Base Year response indicator, all of the estimated coefficients differ statistically from zero. Of the remaining predictors, the most important from the standpoint of the size of the estimated unstandardized coefficient was the indicator of nonresponse to the second follow-up, SFUNR.* This was followed in order by the indicator of nonresponse to the first follow-up, FFUNR; the indicator of personal interview response to the second follow-up, SFUPI; and the indicator of personal interview response to the first follow-up, FFUPI. In every instance nonresponse or personal interview response served to lessen

* In most cases the interpretation of the relative importance of regression coefficients is best accomplished using standardized coefficients since the variability of most scales is highly arbitrary; however, the fact that all of the variables involved were dichotomous makes the unstandardized coefficients more meaningful in this instance from the standpoint of understanding the relative response rates of the various groups involved. This is the position taken with respect to the interpretation of the results of the discriminant analyses presented in the next section of the report as well.

the likelihood of response to the third follow-up, with nonresponse having a markedly larger effect than personal interview response.

The only difference between the third regression equation and the second is that the interaction terms involving mode of response to the first two follow-ups have been added. Inspection of the results for this equation reveals that this overall model is also statistically significant. In addition, the R^2 shows a even smaller but again statistically significant increase over that of the second equation, from .35 to .36. The regression coefficients associated with individual predictors in the model, with the exception of the base year indicator, differ statistically from zero. The indicator of second follow-up nonresponse remains the most important predictor in terms of the size of its estimated regression coefficient. Of the remaining coefficients, the largest were associated with the interaction term indicating nonresponse to both previous follow-ups, FFURN*SFUNR; the indicator of nonresponse to the first follow-up, FFUNR, and the interaction term indicating nonresponse to the first follow-up and personal interview response to the second follow-up, FFUNR*SFUPI. On the whole these results point to the importance of the recency of response together with a history of responding to the follow-up surveys and in the most cooperative manner--by mail.

The final estimated regression equation included all predictors used previously plus the background variables (indicators of sex, ethnicity, high school program, and SES). In general, the addition of the background variables made but a small change in the predictability of third follow-up response. The increase in the R^2 over that of the third regression equation was quite small, less than one tenth of one percent, and did not differ statistically from zero. All of the estimated regression coefficients appearing as statistically significant in the first three regression equations remained statistically significant here as well. As might be expected given the nonsignificant increase in R^2 with the addition of the background variables, only three of the seven regression coefficients for background indicators (mid-level SES, black ethnicity, and white ethnicity) differed statistically from zero. All of these carried positive signs indicating that sample members having mid-level SES, black and white ethnicity were more likely to provide third follow-up responses than the base groups (low SES, and "other" ethnicity); however, the magnitude of all three coefficients was less than .02, making them of less importance in an absolute sense than any of the prior response indicators.

Further inspection of Table 13 reveals that each of the R^2 values for the four regression equations was statistically greater than zero. The changes in R^2 from one equation to the next indicates that, except for the smallest change, all were statistically greater than zero.

D. Discussion of Results

The analyses provide answers to the basic questions that motivated this series of regression analyses. Based on the results presented above, it is clear that a reasonable degree of prediction (about one-third of the variance) of third follow-up response was possible given knowledge of response to the second follow-up. With respect to the other three questions, addition of sets of variables did make statistically significant contributions to the prediction of third follow-up response, with the exception of background variables.

A distinction should be made at this point, however, between statistical and practical significance. Statistical significance actually addresses only whether the added variables should be considered to have non-zero contributions to the prediction of third follow-up response. Practical significance concerns the extent to which that improvement should be considered useful. While this is a more subjective criterion it would seem reasonable to consider the prediction afforded by the second follow-up nonresponse indicator to be of practical significance. Beyond that, the addition of the other variables in the second and third model while making statistically significant improvements in prediction, actually added very little in a practical sense. In fact, the interaction terms added with the third equation increased the R^2 by less than .01--certainly not a substantial increase.

With respect to predicting response to future follow-up surveys involving this or a similar population, it seems reasonable to make predictions using the nonresponse indicator to the immediately preceding follow-up survey alone. Clearly, where one is attempting to predict response to the first survey in a longitudinal series, such information will not be available, and it will be necessary to employ other information, perhaps related to the background of the sample members, to make a prediction of response. Inasmuch as the correlations between the background variables and base year response can speak to this, it would not seem that the quality of this sort of prediction would begin to approach that where prior response information is readily available.

IV. PREDICTING MAIL AND PERSONAL INTERVIEW RESPONSE TO THE THIRD FOLLOW-UP

In predicting the way in which subjects responded to the Third Follow-Up Survey, as distinct from the prediction of response/nonresponse covered in the previous section, the variables involved are the same as those used previously; however, the analytic approach differs somewhat. Because only those who failed to respond by mail were eligible for the personal interview portion of the survey, it was expected that the predictor variables would play somewhat different roles in differentiating each group from the rest. Variables associated with subjects being opposed to respond or being impossible to locate should have the same relative importance vis á vis each group of respondents. On the other hand, variables associated with subjects being somewhat reluctant to respond or less easily located should have different patterns of results for each respondent group. In the following subsections the methodological approach is explained, and the results are presented and discussed.

A. General Approach

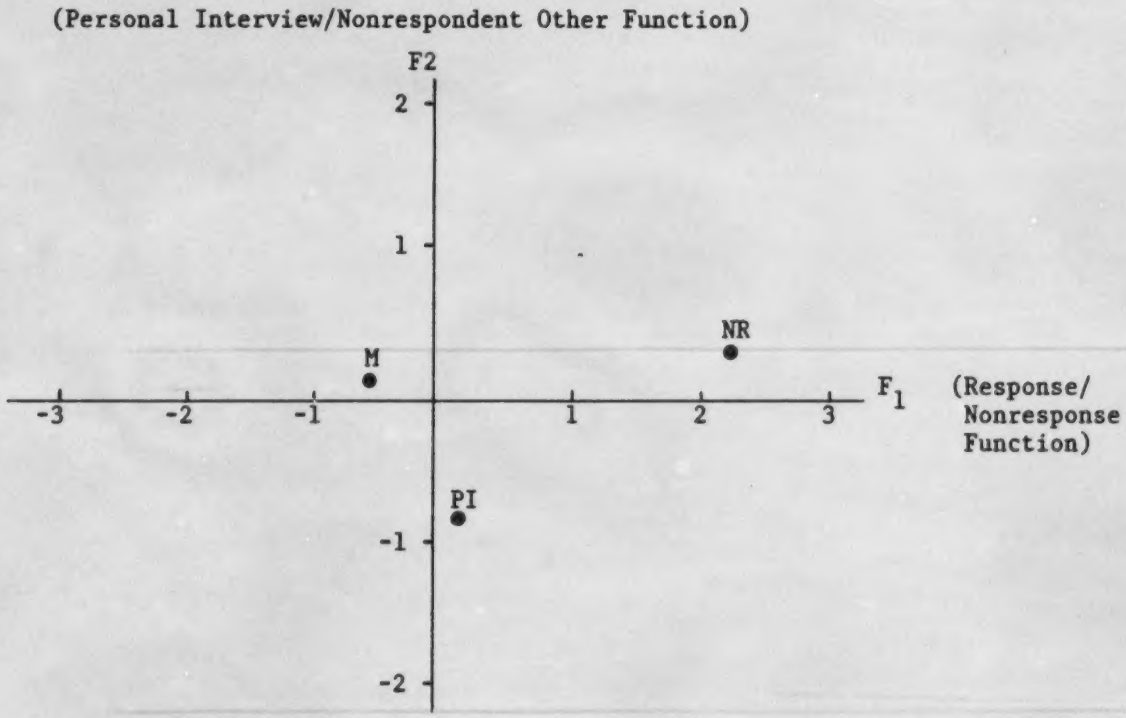
The variable of interest in this section of the report is actually a three-valued categorical variable (indicating mail response, personal interview response, or nonresponse to the survey), which is not amenable to prediction through the use of regression analysis. Instead, discriminant function analysis, an analytic technique allowing prediction of group membership for more than two groups, was used. Where regression analysis yields a single set of coefficients that, when applied to their associated variables, produces the best estimate of the dependent variable (in a least squares sense), discriminant function analysis in the three group case results in two sets of coefficients that define orthogonal dimensions which serve to maximally discriminate the members of the three groups (corresponding to the three mode-of-response categories). Where regression analysis provides a multiple correlation coefficient that reflects the degree to which the estimated regression equation predicts the dependent variable, discriminant analysis provides similar information in the form of a canonical correlation coefficient for each function.*

* A canonical correlation coefficient is a correlation coefficient between optimal linear combinations of two sets of variables. In this case the predictors define one set of variables; the other is based on the response mode categories.

B. Results of the Discriminant Function Analyses

As with the regression analyses used to predict third follow-up response, the discriminant function analyses were carried out in a stagewise fashion with the indicators of base year response and response to the first two follow-ups entered first followed by the addition of the mode of response interaction terms, and, finally, the background variables. To conserve space, only the full model summary results are presented and discussed; more complete results of the analyses are provided in Appendix B.

The centroids for the three groups on each of the two discriminant functions are plotted in Figure 1 to facilitate interpretation of the results.



Group	F_1	F_2
M = Mail respondent	-0.37	0.15
PI = Personal interview	0.20	-0.76
NR = Nonrespondent	2.21	0.29

Figure 1.--Plot of response mode group centroids on the two discriminant functions

Inspection of this plot reveals that the primary role of the first function is to discriminate nonrespondents from respondents while that of the second function is to discriminate those who responded by personal interview from the rest. Clearly, the most salient dimension here was that discriminating the respondents from the nonrespondents. It is also clear, however, that the second function played an important role in discriminating those who responded by mail from those who responded by personal interview. If the three groups could have been discriminated using a single function, their centroids would have come much closer to lying on a single straight line.

Table 14 provides another summary of the full-model analysis. The eigenvalues related to the discriminant functions at the bottom of Table 14 indicate that the majority of the group discrimination taking place was accomplished by the first function with much less contributed by the second. The canonical correlations provide information similar to that yielded by the multiple correlation coefficient in regression analysis. Furthermore, since the functions are orthogonal, their squared values are additive. Thus, the first function can be thought of as having accounting for roughly 38 percent of the between group variance while the second function accounted for an additional 11 percent.

The standardized and unstandardized discriminant function coefficients are also set forth in Table 14 for each of the functions. The variables contributing most to the discrimination of the nonrespondents from the respondents in the order of their unstandardized coefficients are the indicators of second follow-up nonresponse, the interaction term indicating nonresponse to both earlier follow-ups, and the indicator of nonresponse to the first follow-up. All of these had positive signs indicating the high probability of nonresponse for sample members who responded to neither of the first two follow-ups. For the function discriminating those who responded by mail from those who responded by personal interview the most important variables are the indicators of nonresponse to both prior follow-ups, personal interview response to the second follow-up, and nonresponse to the first coupled with personal interview response to the second.

Where the three most important variables with respect to the first function each carried positive weights (indicating a greater likelihood of nonresponse for individuals in the category assigned a value of 1), the situation

Table 14.--Standardized and unstandardized discriminant function coefficients

Variable	Unstandardized coefficients		Standardized coefficients	
	F_1	F_2	F_1	F_2
BYR	0.04841	-0.06660	0.02091	-0.02877
FFUPI	0.39379	-0.61012	0.17701	-0.27425
FFUNR	1.02378	-0.59956	0.20286	-0.11880
SFUPI	0.69463	-1.32398	0.28974	-0.55225
SFUNR	3.89638	0.29505	0.78259	0.05926
FFUPI*SFUPI	0.30309	-0.48270	0.10211	-0.16262
FFUPI*SFUNR	0.51321	0.98260	0.08065	0.15442
FFUNR*SFUPI	-0.92571	1.01349	-0.12529	0.13717
FFUNR*SFUNR	1.26551	2.32669	0.16379	0.30113
SESM	-0.05519	-0.07026	-0.02757	-0.03510
SESH	-0.04954	-0.06093	-0.02079	-0.02557
HPMGEN	0.01985	0.00499	0.00952	0.00239
HPMACAD	0.01775	0.06820	0.00866	0.03327
ETHBLK	-0.00015	-0.61916	-0.00005	-0.21079
ETHWHT	-0.14461	0.24449	-0.06020	0.10178
SEX	-0.04530	0.21422	-0.02259	0.10681
CONSTANT	-0.52687	0.20036		
Eigenvalue	.62177	.12501		
Canonical Correlation	.61919	.33335		

for the second function differed somewhat. In this case, second follow-up personal interview response was associated with a coefficient indicating the greater likelihood of personal interview response on the third follow-up. The signs on the coefficients for the other two variables indicate a greater likelihood of response by mail for those who were nonrespondents to both prior follow-ups and those who both failed to response to the first follow-up and responded by personal interview to the second. Another point which should be noted is that the magnitudes of the coefficients associated with the background variables of ethnicity and sex were greater on the second discriminant function than on the first. The pattern of coefficients indicates a greater likelihood for personal interview response to have been obtained from blacks and males than from others. This is consistent with the results observed in the simple descriptive statistics presented earlier.

C. Interpretation of Results

As indicated above, the two discriminant functions served to differentiate the nonrespondents and personal interview respondents from the mail respondents in that order. Of the two functions the strongest by far was the first accounting for more than a third of the between groups variance. In general, the variables involved played virtually the same role with respect to their coefficients on the first discrminant function as they did in the regression analysis used to predict third follow-up response/nonresponse earlier. As a result, the reader is directed to the interpretation of the results of that analysis for a more substantive interpretation.

Given third follow-up response, the factors tending to discriminate those who responded by mail from those who responded by personal interview form a more complex pattern than was the case for the discrimination of the respondents from the nonrespondents. This would appear to be the outcome of both personal and procedural factors. Thus, the fact that the single most important variable on the second discriminant function was the interaction term indicating nonresponse to both prior follow-ups and carried a positive sign simply reflects the procedural decision which excluded those who consistently failed to respond to the follow-up surveys from personal interview activities in the third follow-up. On the other hand, the coefficient asscciated with the indicator of second follow-up nonresponse serves to reaffirm the importance of recency as a construct in predicting behavior with respect to the third follow-up.

Somewhat more perplexing was the coefficient associated with the first follow-up nonresponse and second follow-up personal interview response indicator. This would seem to indicate a combination of personal and procedural forces aimed at making a mail third follow-up response more likely from someone who both failed to respond to the first follow-up and responded by personal interview to the second all other things being equal. This could be interpreted as the outcome of finding individuals on the second follow-up who were missed on the first and obtaining an address at which the mailed out Third Follow-Up Questionnaire would be received.

V. PREDICTING THE QUALITY OF RESPONSE TO THE THIRD FOLLOW-UP

Previous sections have shown that it is possible to predict, to some extent, response to the Third Follow-Up Survey and to a lesser extent, whether or not that response was more likely to have been obtained by mail or through the personal interview effort. An additional issue that needs to be addressed, from the standpoint of statistical prediction, concerns the quality of the responses that were obtained.

A. General Approach

In section II of this report differences in data quality were observed between those responses obtained by mail and those collected through the personal interview, with the personal interview information generally subject to fewer problems. Due to the disproportionate representation of groups defined by prior response history, ethnicity, SES, etc., the relative importance of each of the variables in determining the quality of the responses is not clear. In addition, the availability of similar data quality indices from the First and Second Follow-Up Surveys naturally leads to the question of the extent to which third follow-up data quality can be predicted through the use of these variables.

Since each of the indices of data quality can be thought of as a continuous variable, regression analysis was used to examine these issues and ten separate analyses were carried out, one for each of the third follow-up data quality indicators. The variables used to predict data quality included the third follow-up personal interview indicator, the first and second follow-up data quality indicators, the background variables already used in the previous analyses, and the other appropriate response history variables. As with the description results, the actual set of cases involved in these analyses was somewhat smaller than those used for the previous analyses. Conditioning on the availability of all of the variables to be used as predictors reduced the number of cases analyzed to 18,954. Since the analyses were applicable only to respondents, nonresponse indicators provide no additional information and, thus, were not used. The names and descriptions of the variables used as predictors here which have not appeared previously are set forth in Table 15.

Table 15.--Variables used in predicting third follow-up data quality indices not appearing elsewhere

Variable	Explanation
TFUPI	Third follow-up mode of response indicator (0 = mail, 1 = personal interview)
FCSIND	First follow-up consistency index
FORIND	First follow-up out of range index
FRINDX	First follow-up routing index
FCIAIN	First follow-up Section A completion index [General Information]
FCIBIN	First follow-up Section B completion index [Education and Training]
FCICIN	First follow-up Section C completion index [Civilian Work Experience]
FCIDIN	First follow-up Section D completion index [Military Service]
FCIEIN	First follow-up Section E completion index [Information About the Past]
CINDEX1	Second follow-up key question consistency index
CINDEX2	Second follow-up record quality index
SFURIND	Second follow-up routing index
SFUORIND	Second follow-up out of range index
SFUCIAIN	Second follow-up Section A completion index [General Information]
SFUCIBIN	Second follow-up Section B completion index [Education and Training]
SFUCICIN	Second follow-up Section C completion index [Work Experience]
SFUCIDIN	Second follow-up Section D completion index [Family Status]

Table 15.--Variables used in predicting third follow-up data quality indices not appearing elsewhere--Continued

Variable	Explanation
SFUCIEIN	Second follow-up Section E completion index [Military Service]
SFUCIFIN	Second follow-up Section F completion index [Activities and Opinions]

B. Results

The results of each of the ten regression analyses are contained in Appendix C of this report. Two sets of data have been drawn from those results to serve as overall summaries. The first set consists of the squared values of the final multiple correlation coefficients for the prediction of each of the dependent variables from all of the independent variables. These are presented at the bottom of Table 16 along with their associated F statistics. Inspection of these results reveal that all of the estimated regression equations were statistically significant but predicted only a very small proportion of the variance in the data quality indicators. The lowest R^2 value (.0056) was associated with the third follow-up out of range index while the largest (.1071) was for the third follow-up routing index.

While none of these regression equations would appear to be particulary useful from the standpoint of identifying specific subgroups for whom special data collection procedures might be warranted, the consistency of relationships existing between the predictors and the dependent variables was examined. Table 16 contains the second set of information drawn from the appendix and consists of a summary of the results of the regression analyses, displaying the signs of the statistically significant regression coefficients for each of the ten data quality indicators. The predictors are arranged along the left while the dependent variables appear across the top. Thus, the pattern of statistically significant results for a regression equation appear in a given column of the table. One important factor that should be kept in mind is that, for the purpose of interpreting this information, all of the data quality indicators have been rescaled so that small values indicate poor data quality and large values good data quality.

With this in mind some obvious patterns emerge from the results in the table if we ignore those for the poorest predicted variable, the out-of-range index. The first point that should be noted is that the indicator of personal interview response to the third follow-up had a statistically significant coefficient for every dependent variable. With the exception of the coefficients related to the prediction of TCIAIND and TCIFIND (which were negative), the data collected through personal interviews would appear to have been better than that collected by mail even when controlling for the remaining background, response history and prior follow-up data quality variables. Coincidentally, inspection of the standardized regression coefficients

Table 16.--Signs of statistically significant regression coefficients associated with variables used to predict third follow-up data quality indices and final R^2 and F values associated with the regression equations

	TFUORIND	TFURIND	TCSIND1	TCSIND2	TCIAIND	TCIBIND	TCICIND	TCIDIND	TCIEIND	TCIFIND
BYR										
FFUPI		-		-		-			-	
SFUPI		-	-	-		-	-	-	-	-
FFUPI*SFUPI										
TFUPI	+	+	+	+	-	+	+	+	+	-
SESH			+				-	+		
SESH	+						-	+		
HPHGEN	+								-	
HPHACAD	+		-		+	+		+	+	+
ETHBLK			-	-		-	-		-	
ETHWHT	-		+	+	+	+	+	+		+
SEX		+	+				+	+	-	+
FCSIND	+		+	+			+	+	+	
FORIND										
FRINDX						+			+	
FCIAIN	+		+				-			
FCIBIN						-				
FCICIN	-		+			+	+	+		
FCIDIN			+					+		
FCIEIN		+	+		+	+	+		+	+
CINDEX1		+	+	+	+	+	+	+		
CINDEX2			+				+			
SFURIND		+	+	+	+	+	+	+	+	
SFUORIND	+					+				
SFUCIAIN		-	-	-	-	-	-	-	-	
SFUCIBIN		+			+		+			-
SFUCICIN		-	-			+		-		
SFUCIDIN		+	+	+	+	+	+	+	+	+
SFUCIEIN		+	+				+	+		
SFUCIFIN									-	+
R^2	.0056	.1071	.0756	.0315	.0451	.0472	.0427	.0286	.0661	.0291
F*	3.52	75.64	51.59	20.51	29.82	31.25	28.16	18.60	44.63	18.89

* An F value with 30 and 18923 degrees of freedom larger than 1.46 is statistically significant at $\alpha = .05$.

presented in Appendix C indicates that the best predictor of data quality was the way it was collected.

Where statistically significant coefficients were associated with the variables indicating personal interview responses to prior follow-ups (FFUPI and SFUPI), their sign indicates that poorer third follow-up data came from those with a history of response through personal interview. Thus it would appear that, while responding through personal interview virtually guarantees better quality data at the time, it indicates that data of poorer quality are likely to be obtained in the future from those who have responded by personal interview in the past.

The relationships involving the background variables and the third follow-up data quality indicators would seem to be as one might expect. The statistically significant regression coefficients associated with sex, ethnicity and high school program variables indicate that the quality of the responses obtained from those who were involved in an academic high school program, who were white, or who were female was better than that from those not having such characteristics.

Turning now to the results associated with the first and second follow-up data quality indices it seems clear that the second follow-up indicators were more likely to have statistically significant regression coefficients than were the first follow-up indicators. Ignoring the results involving TFUORIND, there were 72 estimated regression coefficients involving the first follow-up indicators of which 24 of them (33 percent) were statistically significant. With respect to the second follow-up indicators, there were 90 coefficients of which 50 (56 percent) were statistically significant. Only two first follow-up indicators had regression coefficients that were statistically significant in the prediction of more than half of the third follow-up indicators. These were the indicators of overall completeness, FCSIND, and completeness of the last section of the questionnaire, FCIEIN. In both instances more complete first follow-up data was associated with better third follow-up data.

Considering the second follow-up data quality indicators, four of them had regression coefficients that were statistically significant for more than half of the dependent variables. These include the consistency index, CINDEX1; the routing index, SFURIND; and the Section A and D completeness indices, SFUCIAIN and SFUCIDIN. The general direction of the results, once again, pointed toward better third follow-up data coming from those with better

second follow-up data. The only exception to this was for SFUCIAIN where the results were reversed.

The final point that should be noted concerns the regression coefficients for the Second Follow-Up completeness indicators involved in the prediction of their third follow-up topical counterparts (see Table 8 and Table 15). Of the six regression coefficients that fell into this category, all were statistically significant including those for SFUCIBIN, SFUCICIN, SFUCIEIN, and SFUCIFIN. Only for the regression coefficient associated with SFUCIAIN was the direction of the relationship counterintuitive. Thus, the quality of the data provided under a specific topical area on the second follow-up was generally indicative of the quality of the data in the parallel section on the third.

C. Discussion of Results

The results of the regression equations predicting the third follow-up data quality indices suggest several conclusions. The first is that the types of errors reflected in the various indices come quite close to the traditional psychometric conception of measurement error, that of a random process (at least in a practical sense). Given the proportion of the variance accounted for by the regression equations, it would be difficult to suggest or justify any special approach that might be taken to improve the quality of data collected in the future.

On the other hand, there were discernable and meaningful patterns among those regression coefficents which were statistically significant. As was noted in the descriptive tables appearing in section II and confirmed here, the data collected through personal interview was generally of a higher quality with respect to completeness and consistency. This was the single most important predictor of the third follow-up data quality indicators. At the same time, personal interview data collection was far more costly than mail data collection. Furthermore, prior response to the other follow-up surveys through personal interview and especially the second follow-up was predictive of poorer third follow-up data quality even with the simultaneous control for the background variables included here. Since those who responded previously by mail had to work through the follow-up questionnaires on their own, one might interpret this set of results as indicative of practice effects operating across the various follow-up surveys. What would be more likely, given the time span involved and the fact that individual selection took place with

respect to whether sample members responded by mail or personal interview, is that some sizeable part of not responding by mail reflects individual disinterest in the survey that translates into both unwillingness to respond without considerable prompting and poor attention to the rather complex demands posed by the questionnaire.

As with the prediction of third follow-up response, the more proximal and directly relevant information would seem to have been more highly related to the quality of the third follow-up data. The background variables and the first follow-up data quality indicators were more weakly related to third follow-up data quality than were the second follow-uUp variables. Furthermore, among the second follow-up completeness indices a pattern of results emerged indicating that the provision of complete or incomplete response to a given topical section of the questionnaire persisted through the second follow-up. Thus, while a great deal of the variability in the various indices was at the level of random error, weak meaningful patterns were detected.

VI. THE EFFECTS OF THE INCENTIVE ON THE THIRD FOLLOW-UP

Two interrelated outcomes from the use of an incentive are of some importance from the standpoint of executing a study involving substantial data collection efforts of the sort inherent in the NLS. The first outcome of interest is the effect of the incentive on the rate of mail response to the Third Follow-Up Survey. The second outcome is the effect of the incentive on the cost of the data collection effort. Two elements of cost savings are considered here: those due to a decrease in the inherently more expensive personal interview effort and those due to the overall acceleration of the mail response schedule. The first of these is quite obvious; the second results from the follow-up activities targeted at those not responding to the initial mailings by a specified date. To the extent that returns are obtained early, the later prompting activities will be reduced in scope thereby resulting in cost savings.

As indicated earlier, a direct assessment of the effects of the incentive is precluded since an incentive check was sent to everyone with the first mailout of the Third Follow-Up Questionnaire. Nonetheless, the experimental study carried out as a part of the field test of the Fourth Follow-Up Survey does provide information that can be used to estimate what the third follow-up mail response rate might have been without the use of the incentive and how data collection costs might have been affected, assuming the use of the same data collection procedures.* Each of these topics is considered below following a brief summary of the results of the experimental study.

A. Results of the Experimental Evaluation of the Incentive Check†

The experimental study on the effects of an incentive check on mail response rates to the fourth follow-up field test demonstrated that those who received an incentive check responded sooner and at an overall higher rate than did those who were not sent such a check. Table 17 summarizes both of

---

* Further details on the methodology and results of this study are available in King (1979).

† Since the NLS has employed a strategy aimed at personal interviews with virtually all nonrespondents to the mail portion of the effort, the rate of response to the mailed survey is obviously the most salient element here.

these findings. This table indicates that the difference in the response rate for those sent the incentive as compared to those not sent the incentive tended to grow over time. Two weeks after the initial mailout their difference was only 4 percentage points; 20 weeks afterward the magnitude of this difference had grown to over 14 percentage points. Interestingly enough, however, the ratio of the response rates for the no incentive group to that for the group receiving the incentive is relatively stable over time. The value of this ratio is .803 at the two week point and slowly declines to .765 by week 20. Since this ratio is far more stable over time than the difference between the two groups in their response rates, subsequent calculations make use of the final ratio of .765 in the determination of mail response rates and the average value of .779 from the first nine weeks to determine the extent to which mail response might have been retarded had no incentive been used in connection with the third follow-up.

Table 17.--Percentage of response over time

Weeks after initial mailout	Cumulative percent of response		No incentive response rate ÷ incentive response rate
	Incentive group	No incentive group	
2	20.8	16.7	.803
3	32.1	25.8	.804
4	39.3	31.2	.794
5	44.5	34.0	.764
6	48.0	36.4	.758
7	52.5	40.6	.773
8	55.3	42.5	.769
9	56.7	43.3	.764
⋮	⋮	⋮	⋮
20	60.4	46.2	.765

NOTE.--From King, D. A. "The Effects of Incentives on Response Rates on the National Longitudinal Study of Educational Effects." Research Triangle Park, North Carolina: Research Triangle Institute, June 1979.

B. Mail Response Rate Implications for the Third Follow-Up Survey

As mentioned earlier, the final mail response rate to the Third Follow-Up Survey was 73.8 percent. Assuming that the mail response rate would have been .765 of this in the absence of the incentive checks, the estimated final mail response rate would have been 56.5 percent. This figure is so much lower than that actually obtained (and implies such a large effect of the incentive on response) that one might be inclined to question its reasonableness; however, the response rates to the previous surveys and pilot studies provide support for believing that a similarly low rate of mail response would have been obtained on the third follow-up had no incentive been used. The relevant response rates are presented in Table 18.

Table 18.--Response rates to previous pilot studies and follow-up surveys

Follow-up	Rate of response to pilot study*	Mail rate of response to follow-up study**
1	64	62
2	62	68
3	54	56 (estimate)

* Drawn from King (1979).

** Drawn from King, & Thorne (1977).

The table indicates that response rates obtained from the pilot studies closely parallel the mail response rates to the follow-up surveys. The estimated third follow-up mail response rate of roughly 56 percent fits reasonably well with the pattern of these statistics. Since the pilot study preceeding the third follow-up did not make use of a monetary incentive, it would appear that the predicted sharp drop in the third follow-up mail response rate is a reasonable expectation.

C. Cost Implications for the Third Follow-Up Survey

Assuming that the mail response rates that would have prevailed on the Third Follow-Up Survey had no incentive been used would have been .779 of those actually obtained during the mail prompting activities and .765 of the final mail response rate, it is possible to estimate the cost savings accruing from the use of the incentive check. The processes involved in this estimation are displayed in Figures 2 and 3 for the mail and field costs respectively. Figure 4 represents the final reconciliation of these two cost savings and an estimate of the actual costs associated with making the incentive available. The results of this reconciliation is an estimate of the net savings to the Third Follow-Up Survey of over $316,000 attributable to the use of the incentive.

Activity (unit cost)	Estimated number of actions without incentive	Actual number of actions for the third follow-up*	Net decrease	Net cost savings
Second questionnaire mailout (1.15)	14,153	11,982	2,171	$ 2,497
First prompting mailgram (1.37)	11,615	8,724	2,891	3,961
Third questionnaire mailout (1.15)	9,813	6,410	3,403	3,913
Second prompting postcard (.20)	9,235	5,668	3,567	713
Final mailgram (1.39)	8,960	5,315	3,645	5,067
Estimated total cost savings				$16,151

* Drawn from King, D. A., & Thorne, N. R. "National longitudinal study: Data collection activities for the third follow-up: July 1976 through June 1977. Final report." Research Triangle Park, North Carolina: Research Triangle Institute, 1977.

Figure 2.--Computation of cost savings in mail effects due to the incentive

	Actual third follow-up experience	Estimated experience without incentive
Number of eligible cases	21,807	21,807
Mail returns	- 16,102	- 12,318
Outstanding cases	5,705	9,489
Deletions from field efforts	- 645	- 645
Total cases assigned for personal interviews	5,060	8,844
Cost/case	92.94	92.94
Total cost	407,283	821,973
Estimated total cost savings with the incentive	$351,690	

Figure 3.--Computation of cost savings in field interviews due to the incentive

Cost savings from mail effort	$ 16,151
Cost savings from field effort	351,690
Total savings	$ 367,841
Less cost of checks mailed (@ $.10 each)	2,181
	$ 365,660
Less cost of checks cashed (@ $3.00 each)	48,930
Net savings	$ 316,730

Figure 4.--Computation of total net savings due to the incentive

VII. SUMMARY OF FINDINGS

A. Third Follow-Up Response

There was relatively little between group variability in response rates to the Third Follow-Up Survey when considering groups defined by the background variables of aptitude, SES, ethnicity, sex or high school curriculum. The only exceptions to this involved subjects classified as "unknown" reflecting persistent nonresponse to the survey instruments. Pronounced differences in response rates were observed with respect to groups defined by their pattern of response to prior instruments with the strongest differences associated with the most recent prior survey--the second follow-up.

When the influence of all of the variables was systematically evaluated through a series of multiple regression analyses the single most important predictor of third follow-up response was found to be response to the second follow-up. The use of additional variables measuring background and other information about the consistency and mode of response to the earlier surveys, while making statistically significant improvements in the prediction equation, contributed little to the prediction of Third Follow-Up response over that accomplished by the second follow-up response indicator alone.

B. Mode of Response to the Third Follow-Up

Much stronger differences in the mode of response to the third follow-up were found among the groups defined by the background variables than was the case with the rate of response. White sample members were much more likely to respond by mail than were members of other ethnic groups, especially blacks. Furthermore, aptitude and SES were also positively related to response by mail.

As was the case with third follow-up response rates, the subgroups defined by their response patterns to the earlier survey instruments displayed the greatest variability. Extremely low rates of response by personal interview were observed; however, as was noted, this reflected procedural decisions to exclude those with a history of nonresponse from the personal interview effort. On the other hand, the fact that those with a history of response through personal interview had the highest rates of third follow-up personal interview response was seen as indicative of low individual motivation to participate in

the survey. Thus, both individual and procedural factors were manifest in the data.

From another point of view, it was abundantly clear that the personal interview effort made a tremendous difference in the extent to which various subgroups were underrepresented in the respondent group. Had only a mail data collection effort been used on the third follow-up, whites would have been half again as likely as blacks to have responded. Thanks to the additional personal interview effort, whites were in fact only seven percent more likely to have responded to the third follow-up.

The discriminant function analyses made it clear that, even with these relatively demonstrable individual and procedural effects, nonresponse was easier to predict than personal interview response. Furthermore, the relative importance of the predictors of nonresponse was somewhat easier to understand than that for the predictors of personal interview response. In both instances, however, response to the second follow-up was one of the most important discriminators with the nonresponse indicator most important for prediction of nonresponse and the personal interview indicator most important for predicting personal interview response. This again confirmed the general finding that the most recent and relevant information functioned as the best predictor.

C. The Quality of Third Follow-Up Responses

As might be expected, there were discernable differences in the quality of the data collected by the mail and personal interview efforts. In general, this difference tended to point to better data having been collected through the personal interview than by mail; however, the results for several completeness indices ran counter to this trend. Because of the characteristics of the personal interview respondents together with their presumed lower interest in the survey, this was interpreted as resulting from the interviewers being able to elicit more consistent responses where responses were given but not being able to overcome the respondents reluctance to provide information in all instances.

The results with respect to subgroups defined by the background variables pointed to better data having been provided by those of higher aptitude, higher SES, from academic high school programs, and who were white. Giver the characteristics of the personal interview respondents, it is quite clear that the personal interview process represented a considerable investment of effort

over and above simply soliciting cooperation. Without the direction provided in the personal interviews it would seem that the overall quality of data collected during the third follow-up would have suffered considerably.

Given these results, it was not surprising that previous response history made little difference in the quality of the data among those who responded through the personal interview. Among those who responded by mail, however, better data was obtained from those who responded by mail to the Second and First Follow-Up Surveys. While the regression analyses yielded statistically significant results and meaningful patterns of significant coefficients, the very low degree to which the data quality indices could be predicted ruled out any practical application of the results.

D. Importance of the Incentive

The results of the experimental evaluation of the use of an incentive check carried out as part of the fourth follow-up field test indicated that the overall mail response rate was substantially enhanced and that the rate at which responses were received during the mail follow-up period was accelerated. When these effects were applied to the third follow-up to estimate what the response would have been had no incentive been used, it was found that very large differences would have existed. The net cost savings accruing from the accelerated mail response and overall higher mail response rate were estimated as being over $316,000.

References

Crider, D. M., & Willits, F. K. Respondent retrieval bias in a longitudinal survey. Sociology and Social Research, 1973, 58(1), 56-65.

King, D. A. The effect of incentives on response rates on the national longitudinal survey of educational effects. Research Triangle Park, North Carolina: Research Triangle Institute, 1979.

King, D. A., & Thorne, N. R. National longitudinal study: Data collection activities for the third follow-up: July 1976 through June 1977. Final report, Research Triangle Park, North Carolina: Research Triangle Institute, 1977.

Lenski, G. E., & Leggett, J. C. Caste, class, and difference in the research interview. The American Journal of Sociology, 1960, 65(5), 463-467.

Levinsohn, J. R., Henderson, L. B., Riccobono, J. A., & Moore, R. P. National longitudinal study base-year, first, second, and third follow-up data file users manual. Washington, D.C.: National Center for Education Statistics, 1978.

Moser, C. A., & Kalton, G. Survey methods in social investigation. New York: Basic Books, Inc., 1972.

Nuess, C. F. Differences between persons responding and not responding to a mailed questionnaire. American Sociological Review, 1943, 8(4), 433-438.

Vincent, C. E. Socioeconomic status and familial variables in mail questionnaire responses. The American Journal of Sociology, 1964, 64(6), 647-653.

Appendix A

Standardized Regression Results for the Prediction of Third Follow-Up Response

Appendix A

Standardized regression results for the prediction of third follow-up response

Variable	Equation 1		Equation 2		Equation 3		Equation 4	
	B	F	B	F	B	F	B	F
BYR	-	-	.007	1.73	.007	1.45	.007	1.75
FFUPI	-	-	-.084	197.04[a]	-.063	70.53[a]	-.061	66.19[a]
FFUNR	-	-	-.100	269.88[a]	-.089	35.59[a]	-.088	34.38[a]
SFUPI	-	-	-.097	268.58[a]	-.092	121.11[a]	-.092	118.22[a]
SFUNR	-.569	10546.59[a]	-.543	8116.78[a]	-.452	1642.40[a]	-.451	1631.63[a]
FFUPI*SFUPI	-	-	-	-	-.035	12.97[a]	-.035	13.05[a]
FFUPI*SFUNR	-	-	-	-	-.054	31.57[a]	-.055	32.33[a]
FFUNR*SFUPI	-	-	-	-	.047	17.25[a]	.047	16.63[a]
FFUNR*SFUNR	-	-	-	-	-.112	71.77[a]	-.112	72.32[a]
SEX	-	-	-	-	-	-	.001	.05
SESH	-	-	-	-	-	-	.016	5.60[a]
SESH	-	-	-	-	-	-	.012	2.85
HPHGEN	-	-	-	-	-	-	-.005	.46
HPHACAD	-	-	-	-	-	-	-.007	.94
ETHBLK	-	-	-	-	-	-	.019	5.53[a]
ETHWHT	-	-	-	-	-	-	.019	5.84[a]
R^2(F)	.32340	(10546.59)[a]	.35338	(2411.23)[a]	.36203	(1390.76)[a]	.36245	(783.48)[a]
R^2 change (F)[b]	.32340	(10546.59)[a]	.02998	(188.94)[a]	.00865	(74.77)[a]	.00042	(2.08)[a]

[a] Statistically significant with $\alpha = .01$.

[b] The F value for the R^2 change associated with Equation 1 is simply that of the R^2 value of the model. The others are for the changes in R^2 attributable to the additional variables incorported and were computed using the following formula:

$$F = \frac{(R_g^2 - R_s^2)/(k_g - k_s)}{(1 - R_g^2)/(N-k_g-1)}$$

Where R_g^2 refers to the larger of the two R^2 values used to compute the change in R^2, k_g is the number of predictor variables in the equation with the larger R_g^2 and N is the number of cases used in the analysis (22,067).

Appendix B

Summary of Discriminant Function Analyses for the Prediction of Mail or Personal Interview Response to the Third Follow-Up Survey

Appendix B

Summary of discriminant function analyses for the prediction of mail or personal interview response to the third follow-up survey

Variable	Wilks' lambda	Sig.	Rao's V*	Sig.	Change in V*	Sig.
BYR	0.997	0.00	70.00	0.00	69.82	0.00
FFUPI	0.923	0.00	1851.00	0.00	1781.00	0.00
FFUNR	0.830	0.00	4393.00	0.00	2542.00	0.00
SFUPI	0.781	0.00	5846.00	0.00	1453.00	0.00
SFUNR	0.565	0.00	15560.00	0.00	9715.00	0.00
FFUPI*SFUPI	0.562	0.00	15730.00	0.00	165.40	0.00
FFUPI*SFUNR	0.562	0.00	15750.00	0.00	19.13	0.00
FFUNR*SFUPI	0.557	0.00	16040.00	0.00	290.90	0.00
FFUNR*SFUNR	0.555	0.00	16160.00	0.00	124.00	0.00
SESM	0.554	0.00	16170.00	0.00	10.19	0.01
SESH	0.554	0.00	16200.00	0.00	27.19	0.00
HPMGEN	0.554	0.00	16200.00	0.00	0.56	0.75
HPMACAD	0.554	0.00	16200.00	0.00	1.55	0.46
ETHBLK	0.550	0.00	16400.00	0.00	204.00	0.00
ETHWHT	0.549	0.00	16440.00	0.00	35.83	0.00
SEX	0.548	0.00	16480.00	0.00	37.46	0.00

* These figures are reported to four significant digits.

Appendix C

Regression Results for the Prediction of the Third Follow-Up Data Quality Indicators

Appendix C
Regression results for the prediction of TFUORIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.0047	-.0074	1.02
FFUPI	.0083	.0138	1.97
SFUPI	.0041	.0064	0.31
FFUPI*SFUPI	-.0069	-.0088	0.47
TFUPI	-.0246	-.0353	20.04
SESM	-.0034	-.0062	0.46
SESH	.0008	.0013	0.02
HPMGEN	-.0039	-.0067	0.52
HPMACAD	-.0030	-.0053	0.29
ETHBLK	.0007	.0009	0.01
ETHWHT	-.0096	-.0144	1.74
SEX	-.0114	-.0206	7.71
FCSIND	.0018	.0114	1.37
FORIND	.0004	.0010	0.02
FRINDX	.0003	.0006	0.33
FCIAIN	-.0006	-.0364	8.27
FCIBIN	.0001	.0074	0.52
FCICIN	.0003	.0284	5.49
FCIDIN	-.0003	-.0143	3.62
FCIEIN	.0001	.0094	0.56
CINDEX1	-.0002	-.0006	0.01
CINDEX2	.0006	.0096	1.15
SFURIND	.0005	.0075	0.65
SFUORIND	.0371	.0411	31.81
SFUCIAIN	.0004	.0258	2.20
SFUCIBIN	-.0000	-.0024	0.06
SFUCICIN	.0001	.0128	1.35
SFUCIDIN	-.0004	-.0336	3.36
SFUCIEIN	-.0002	-.0079	1.06
SFUCIFIN	-.0001	-.0109	0.25
Constant	.0989		
R^2	.0056		
F	3.52		

Regression results for the prediction of TFURIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	.0677	.0084	1.46
FFUPI	.3567	.0469	25.25
SFUPI	.5738	.0703	42.50
FFUPI*SFUPI	.1117	.0112	0.85
TFUPI	-1.5977	-.1802	580.92
SESM	-.0832	-.0119	1.87
SESH	-.1531	-.0186	4.10
HPMGEN	-.1431	-.0195	4.92
HPMACAD	-.3744	-.0526	31.70
ETHBLK	.5702	.0548	28.68
ETHWHT	-.4466	-.0527	25.93
SEX	-.0905	-.0129	3.38
FCSIND	.1566	.0780	71.45
FORIND	.0075	.0014	0.04
FRINDX	.0128	.0178	3.21
FCIAIN	.0002	.0009	0.01
FCIBIN	-.0009	-.0078	0.63
FCICIN	-.0028	-.0190	2.75
FCIDIN	.0031	.0106	2.24
FCIEIN	-.0037	-.0254	4.54
CINDEX1	.0810	.0228	9.22
CINDEX2	.0010	.0012	0.02
SFURIND	.1422	.1581	318.88
SFUORIND	-.0224	-.0020	0.08
SFUCIAIN	.0188	.0893	29.29
SFUCIBIN	-.0031	-.0253	7.72
SFUCICIN	.0042	.0321	9.45
SFUCIDIN	-.0115	-.0752	18.73
SFUCIEIN	-.0117	-.0345	22.31
SFUCIFIN	-.0013	-.0083	0.16
Constant	2.5952		
R^2	.1071		
F	75.64		

Regression results for the prediction of TCSIND1

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	.0113	.0037	0.27
FFUPI	.0177	.0061	0.42
SFUPI	.1751	.0563	27.34
FFUPI*SFUPI	-.0075	-.0020	0.03
TFUPI	-.6324	-.1872	605.86
SESM	-.0539	-.0202	5.23
SESH	-.0240	-.0076	0.67
HPMGEN	-.0251	.0090	1.01
HPMACAD	.0947	.0349	13.50
ETHBLK	.2764	.0698	44.89
ETHWHT	-.1634	-.0506	23.10
SEX	-.1385	-.0519	52.65
FCSIND	.0519	.0678	52.20
FORIND	.0008	.0004	0.00
FRINDX	-.0015	-.0056	0.31
FCIAIN	.0027	.0350	8.22
FCIBIN	.0001	.0024	0.06
FCICIN	-.0022	-.0393	11.28
FCIDIN	-.0032	-.0282	15.13
FCIEIN	-.0017	-.0315	6.78
CINDEX1	.0730	.0540	49.95
CINDEX2	.0042	.0132	2.35
SFURIND	.0205	.0600	44.31
SFUORIND	-.0336	-.0077	1.20
SFUCIAIN	.0068	.0853	25.77
SFUCIBIN	-.0004	-.0087	0.89
SFUCICIN	.0011	.0229	4.63
SFUCIDIN	-.0041	-.0699	15.65
SFUCIEIN	-.0069	-.0530	50.88
SFUCIFIN	-.0008	-.0136	0.42
Constant	1.5728		
R^2	.0756		
F	51.59		

Regression results for the prediction of TCSIND2

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	.0226	.0059	0.66
FFUPI	.1102	.0302	9.68
SFUPI	.1857	.0475	17.87
FFUPI*SFUPI	-.0319	-.0067	0.28
TFUPI	-.3541	-.0833	114.52
SESM	-.0487	-.0145	2.57
SESH	-.0465	-.0118	1.52
HPMGEN	-.0086	-.0024	0.07
HPMACAD	-.0363	-.0106	1.20
ETHBLK	.1301	.0261	6.00
ETHWHT	-.1982	-.0488	20.49
SEX	.0432	.0129	3.09
FCSIND	.0386	.0401	17.41
FORIND	-.0178	-.0067	0.88
FRINDX	.0064	.0185	3.21
FCIAIN	-.0003	-.0026	0.04
FCIBIN	.0008	.0138	1.86
FCICIN	-.0014	-.0196	2.69
FCIDIN	.0010	.0068	0.84
FCIEIN	-.0011	-.0153	1.53
CINDEX1	.0270	.0159	4.11
CINDEX2	.0161	.0402	20.88
SFURIND	.0189	.0437	22.50
SFUORIND	.0012	.0002	0.00
SFUCIAIN	.0079	.0784	20.79
SFUCIBIN	-.0006	-.0105	1.23
SFUCICIN	-.0004	-.0067	0.38
SFUCIDIN	-.0053	-.0730	16.31
SFUCIEIN	-.0017	-.0102	1.81
SFUCIFIN	.0006	.0079	0.14
Constant	.8512		
R^2	.0315		
F	20.51		

Regression results for the prediction of TCIAIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.2745	-.0080	1.24
FFUPI	-.4900	-.0151	2.46
SFUPI	-.3026	-.0087	0.61
FFUPI*SFUPI	-.0526	-.0012	0.01
TFUPI	-5.2909	-.1402	329.17
SESM	.2562	.0086	0.92
SESH	.1109	.0032	0.11
HPMGEN	.2416	.0078	0.72
HPMACAD	.8383	.0277	8.21
ETHBLK	-.1055	-.0024	0.05
ETHWHT	.8015	.0222	4.32
SEX	.0467	.0016	0.05
FCSIND	.0034	.0004	0.00
FORIND	.1313	.0056	0.61
FRINDX	-.0458	-.0150	2.14
FCIAIN	-.0061	-.0070	0.31
FCIBIN	.0045	.0088	0.76
FCICIN	.0056	.0091	0.59
FCIDIN	.0024	.0019	0.07
FCIEIN	.0162	.0263	4.56
CINDEX1	-.3051	-.0202	6.77
CINDEX2	-.0592	-.0167	3.65
SFURIND	-.0842	-.0220	5.78
SFUORIND	-.3321	-.0068	0.91
SFUCIAIN	-.0414	-.0464	7.38
SFUCIBIN	.0116	.0219	5.42
SFUCICIN	-.0006	-.0011	0.01
SFUCIDIN	.0596	.0918	26.12
SFUCIEIN	-.0145	-.0100	1.75
SFICIFIN	.0049	.0075	0.12
Constant	93.0615		
R^2	.0451		
F	29.82		

Regression results for the prediction of TCIBIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.3856	-.0093	1.65
FFUPI	-2.3810	-.0603	39.16
SFUPI	-2.7830	-.0657	34.80
FFUPI*SFUPI	.9741	.0188	2.26
TFUPI	1.3522	.0294	14.48
SESM	.6137	.0167	3.54
SESH	.3300	.0077	0.66
HPMGEN	-.0272	-.0007	0.01
HPMACAD	.9018	.0244	6.40
ETHBLK	-1.9775	-.0367	12.01
ETHWHT	1.4840	.0338	9.96
SEX	-.4540	-.0125	2.96
FCSIND	-.1643	-.0158	2.74
FORIND	-.0028	-.0001	0.00
FRINDX	-.0940	-.0253	6.08
FCIAIN	-.0222	-.0209	2.84
FCIBIN	.0132	.0212	4.45
FCICIN	.0253	.0335	7.98
FCIDIN	-.0123	-.0080	1.20
FCIEIN	.0228	.0302	6.05
CINDEX1	-.3489	-.0189	5.96
CINDEX2	-.0724	-.0167	3.68
SFURIND	-.2819	-.0604	43.63
SFUORIND	-.9482	-.0159	4.98
SFUCIAIN	-.0877	-.0804	22.24
SFUCIBIN	.0116	.0180	3.66
SFUCICIN	.0182	.0266	6.08
SFUCIDIN	.0743	.0938	27.30
SFUCIEIN	.0015	.0008	0.01
SFUCIFIN	-.0102	-.0127	0.36
Constant	89.8976		
R^2	.0472		
F	31.25		

Regression results for the prediction of TCICIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.4981	-.0096	1.77
FFUPI	-.3151	-.0064	0.44
SFUPI	-2.1543	-.0408	13.39
FFUPI*SFUPI	-.2290	-.0036	0.08
TFUPI	5.9162	.1031	177.99
SESM	-.8019	-.0177	3.88
SESH	-1.5405	-.0289	9.29
HPMGEN	0.4945	.0104	1.31
HPMACAD	-.2427	-.0053	0.30
ETHBLK	-2.5564	-.0380	12.89
ETHWHT	2.5276	.0462	18.56
SEX	1.2891	.0285	15.31
FCISIND	-.4374	-.0337	12.45
FORIND	.1918	.0054	0.57
FRINDX	-.0694	-.0150	2.13
FCIAIN	-.0440	-.0333	7.18
FCIBIN	.0062	.0079	0.62
FCICIN	.0269	.0286	5.78
FCIDIN	.0234	.0123	2.78
FCIEIN	.0311	.0332	7.24
CINDEX1	-.5967	-.0260	11.19
CINDEX2	-.1383	-.0256	8.61
SFURIND	-.3007	-.0518	31.88
SFUORIND	-.2173	-.0029	0.17
SFUCIAIN	-.1209	-.0890	27.13
SFUCIBIN	.0489	.0610	41.76
SFUCICIN	.0056	.0065	0.37
SFUCIDIN	.0723	.0733	16.61
SFUCIEIN	.0481	.0219	8.39
SFUCIFIN	-.0077	-.0077	0.13
Constant	74.7064		
R^2	.0427		
F	28.16		

Regression results for the prediction of TCIDIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.1591	-.0061	0.69
FFUPI	-.3115	-.0125	1.65
SFUPI	-1.1854	-.0444	15.56
FFUPI*SFUPI	.4876	.0149	1.39
TFUPI	1.2745	.0439	31.71
SESM	.5022	.0219	5.85
SESH	.6510	.0241	6.36
HPMGEN	.2708	.0113	1.51
HPMACAD	.6336	.0272	7.79
ETHBLK	-.6213	-.0182	2.92
ETHWHT	.8120	.0293	7.35
SEX	1.1974	.0522	50.72
FCSIND	-.2211	-.0336	12.22
FORIND	-.1066	-.0059	0.67
FRINDX	.0071	.0030	0.08
FCIAIN	-.0065	-.0097	0.61
FCIBIN	-.0062	-.0158	2.41
FCICIN	.0169	.0356	8.81
FCIDIN	.0364	.0377	25.86
FCIEIN	.0035	.0064	0.27
CINDEX1	-.3475	-.0299	14.57
CINDEX2	-.0026	-.0010	0.01
SFURIND	-.1064	-.0362	15.33
SFUORIND	.2774	.0074	1.05
SFUCIAIN	-.0310	-.0451	6.87
SFUCIBIN	-.0034	-.0085	0.79
SFUCICIN	-.0093	-.0215	3.88
SFUCIDIN	.0235	.0470	6.72
SFUCIEIN	.0761	.0683	80.53
SFUCIFIN	.0095	.0186	0.75
Constant	86.3283		
R^2	.0287		
F	18.60		

Regression results for the prediction of TCIEIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.4325	-.0120	2.82
FFUPI	-1.3819	-.0404	17.89
SFUPI	-2.9410	-.0800	52.70
FFUPI*SFUPI	.8200	.0182	2.17
TFUPI	6.2488	.1566	419.39
SESM	.1715	.0054	0.38
SESH	.1179	.0032	0.12
HPMGEN	.9517	.0288	10.27
HPMACAD	1.8464	.0576	36.39
ETHBLK	-.9841	-.0210	4.03
ETHWHT	-.1233	-.0032	0.09
SEX	-.5443	-.0173	5.77
FCISIND	-.3447	-.0381	16.33
FORIND	-.1743	-.0070	0.99
FRINDX	-.0701	-.0217	4.58
FCIAIN	.0140	.0152	1.53
FCIBIN	.0018	.0034	0.12
FCICIN	.0133	.0203	2.99
FCIDIN	-.0007	-.0005	0.00
FCIEIN	.0179	.0274	5.07
CINDEX1	-.1246	-.0078	1.03
CINDEX2	.0225	.0060	0.48
SFURIND	-.1978	-.0489	29.13
SFUORIND	-.4908	-.0095	1.81
SFUCIAIN	-.1480	-.1565	85.98
SFUCIBIN	-.0023	-.0041	0.20
SFUCICIN	.0121	.0204	3.65
SFUCIDIN	.1942	.2825	252.87
SFUCIEIN	-.0096	-.0063	0.70
SFUCIFIN	-.0526	-.0755	12.79
Constant	84.7920		
R^2	.0661		
F	44.63		

Regression results for the prediction of TCIFIND

Variable	Unstandardized coefficient	Standardized coefficient	F
BYR	-.0757	-.0021	0.08
FFUPI	-.4488	-.0132	1.86
SFUPI	-1.4251	-.0392	12.16
FFUPI*SFUPI	.6171	.0139	1.21
TFUPI	-2.4006	-.0608	60.83
SESM	.2771	.0089	0.96
SESH	-.0995	-.0027	0.08
HPMGEN	.4389	.0134	2.14
HPMACAD	.6883	.0217	4.97
ETHBLK	.1337	.0029	0.07
ETHWHT	1.1409	.0302	7.85
SEX	.9727	.0312	18.10
FCSIND	.0160	.0018	0.04
FORIND	-.0293	-.0012	0.03
FRINDX	-.0421	-.0132	1.62
FCIAIN	-.0155	-.0170	1.85
FCIBIN	.0048	.0090	0.78
FCICIN	.0019	.0030	0.06
FCIDIN	-.0054	-.0041	0.30
FCIEIN	.0342	.0529	18.16
CINDEX1	-.1616	-.0102	1.70
CINDEX2	-.0589	-.0158	3.24
SFURIND	-.1141	-.0285	9.53
SFUORIND	-.4187	-.0082	1.30
SFUCIAIN	-.0968	-.1034	36.13
SFUCIBIN	.0047	.0085	0.80
SFUCICIN	.0037	.0063	0.33
SFUCIDIN	.0674	.0990	29.90
SFUCIEIN	-.0200	-.0132	3.00
SFUCIFIN	.0339	.0491	5.20
Constant	87.8244		
R^2	.0291		
F	18.89		